MODELING OF HETEROJUNCTION BIPOLAR TRANSISTORS

By

CHETAN D. PARIKH

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1991

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MODELING OF HETEROJUNCTION BIPOLAR TRANSISTORS

By

Chetan D. Parikh

May 1992

Chairman: Prof. Fredrik A. Lindholm
Major Department: Electrical Engineering

Analytical models are derived which take into account the various
phenomena occurring in heterojunction bipolar transistors (HBTs). In
particular (i) a charge-control model is derived which is applicable to
abrupt and graded, as well as single and double heterojunction
transistors, and for arbitrary doping profiles in the base; (ii) space
charge region recombination in abrupt and graded heterojunctions is
modeled, and the model is incorporated into the charge-control model;
(iii) a model for quasi-saturation in the collector of a GaAs HBT is
derived, which includes the effects of the drift velocity versus
electric field characteristics of electrons in GaAs; high current
effects in the collector of a double heterojunction transistor are
modeled; (iv) non-quasi-static effects in the base of an AlGaAs/GaAs HBT
are modeled. Comparison with experimental and numerical data, wherever
feasible, show excellent agreement with the models proposed.

The difficulties in modeling velocity overshoot effects and forward biased junction capacitance, and fabrication and dc measurements on AlGaAs/GaAs heterojunctions are described.

Detailed literature reviews, and many tutorial overviews, of various aspects of HBT operation are presented.

CHAPTER 1
INTRODUCTION

This dissertation is concerned with understanding and modeling the behavior of heterojunction bipolar transistors (HBTs).

A heterojunction bipolar transistor is a transistor in which the emitter-base junction (and sometimes the base-collector junction also) is a heterojunction, with the emitter having a larger energy gap than the base. This transistor structure was first conceived by Shockley [Sho51]; he realized that having a wide-gap emitter will reduce minority carrier injection into the emitter, and thus nearly ideal injection efficiency would be achieved irrespective of the base and emitter doping densities. This would lead to an increased current gain. Kroemer [Kro59] reported the first analytical treatment of this structure. Over the years, researchers have realized that such independent control of the injection efficiency allows heavy base and light emitter doping densities, leading to reduced base resistance and emitter-base junction capacitance. These reductions result in higher device speeds and superior high frequency performance.

Developments in fabrication technologies, especially molecular beam epitaxy and metal organic chemical vapor deposition, have improved the quality of HBTs. A brief history of the developments in HBT technology is given in Sec. 4.8. Initially AlGaAs/GaAs and InGaAs/InP transistors, and lately Si/SiGe transistors, have been demonstrated as viable options for microwave and very high speed applications. A GaAs

1

HBT with a cutoff frequency of 90 GHz [Nag89] and a 32-bit CPU with 12,400 gates, made with GaAs HBTs [Eva87] have been reported.

With GaAs and InP HBTs being used in microwave applications, and with very large scale integrated circuits (VLSIs) based on HBTs being fabricated [Ede82, Hall87, Asb89a Asb89b, Yua89], there is a concurrent need for analytical models that can be used in circuit simulators. The work already done in deriving such models is reviewed in chapters 3 and 4. Much remains to be done. This dissertation has been an endeavor at making a contribution to the existing literature on HBT modeling. The following is an overview of this dissertation.

A new charge-control model is derived for the HBT (chapter 5). The model is applicable to single and double heterojunction transistors, with abrupt or graded heterojunctions. It includes effects of carrier flow mechanisms such as thermionic emission and tunneling across the heterojunction; and it is applicable for arbitrary doping density profiles in the base. Space charge region recombination current (I_{SCR}) is modeled, and the model is incorporated into the charge-control model (chapter 6). The model takes into account the variations in the energy gap across the junction. An interesting and unique interrelationship is discovered between I_{SCR} and the collector current. High current density phenomena in the collector of N/p/n GaAs transistors are then modeled (chapter 7); the model takes into account the drift velocity versus electric field characteristics of electrons in GaAs. Some consideration is also given to effects at high current densities in double heterojunction bipolar transistors (i.e. transistors in which both the

emitter-base and the base-collector junctions are heterojunctions). A non-quasi-static model is derived for HBTs (chapter 8) that considers non-quasi-static effects in the quasi-neutral base of the transistor.

Chapters 2 through 4 are part review and part tutorial in nature. In chapter 2, a brief description of the problem of energy band lineups at a heterojunction interface is presented, with discussion of some issues that have caught this author's attention. In chapter 3, a review of the literature on current flow mechanisms across heterojunctions is given, followed by short tutorials on thermionic emission and tunneling -- two current mechanisms that are significant in heterojunction carrier transport. Chapter 4 reviews four different aspects of HBT literature: the analytical models proposed for HBTs (sections 4.2 through 4.5); studies on hot electron transport in the base of HBTs (Sec. 4.6); studies on velocity overshoot and its exploitation in the base-collector space charge region of HBTs (Sec. 4.7); and the developments in the fabrication technologies of compound semiconductor HBTs (Sec. 4.8).

Then lastly in chapter 9, three issues are described that were looked at during the course of this work, but that did not yield fruitful results. The dissertation is concluded in chapter 10.

Wherever possible, experimental and numerical simulation data have been taken from published literature to compare with the analytical models proposed. The corroboration of the proposed models has not been comprehensive; but this has been unavoidable. The difficulties in fabricating devices are discussed in Sec. 9.1. Numerical simulators have not been available.

Throughout the work the AlGaAs/GaAs system has been used as the example, but the models derived are applicable to any heterostructure system (the only exception is the model of chapter 7, which is specific to GaAs).

Some good general references on heterojunctions are the works by Milnes [Miln72, Mil86, Mil87] and Shur [Shu87].

A knowledgeable reader may find that the review chapters (2 through 4) are a bit long and do not contain much new information. Such a reader is requested to cursorily skim through these chapters and concentrate on chapters 5 through 9, which deal with the main contributions of this dissertation. The reason for including elaborate review-cum-tutorial chapters are the following. The information contained in these chapters was gathered as a necessary prelude to the dissertation, and considerable amount of time was spent in this process. It is therefore hoped that a future student of this subject would benefit from this effort, and would not need to spend as much time gathering the same information as this author had to. Also, there came up issues while reviewing the literature that both fascinated and intrigued this author (for example, thermionic emission, velocity overshoot, and others); these issues are presented as short tutorials both to record the author's understanding and to aid other students of the subject. But ultimately perhaps, a dissertation is defined as a "formal and systematic exposition of a subject," and thus should not merely contain the original contributions of the author.

CHAPTER 2
ENERGY BAND LINE-UPS AT HETEROJUNCTION INTERFACES

2.1 Introduction

A semiconductor heterojunction is, by definition, a junction between two dissimilar semiconductor materials, having dissimilar band structures and therefore unequal energy gaps. The study of the behavior of a heterojunction, therefore, begins with an understanding of how the energy bands get aligned at the hetero-interface.

Much research has been done on the nature of the band diagram at the heterojunction interface. There have been three major approaches to determining the band diagram. One approach starts with the Schrodinger equation, solves it in the two semiconductors and matches boundary conditions at the interface [Fre76, Har77, Fre77, Tej77, Schul78, Ihm79, Kan80, Kro83a, Ter84, Hei88]. The solution is, of course, approximate, and relies on some self-consistent method or another. The second approach uses the so-called "electron affinity rule" [And62], which we will look at in Sec. 2.3. The third approach infers the band line-up from experimental data [How73, Mar84, Kro84, Kro85a, Bate85, Kat85, Oku85, Wat85, Wang85, Lin85, Bat86, Wan86, Hei87].

The determination of band line-ups is a very complex problem and we will not address it in this work. In this chapter, we will take a qualitative look at the concept of energy band line-ups (Sec. 2.2); we will describe the electron affinity rule (Sec. 2.3), and discuss the "local vacuum level."

2.2 The Concept of Energy Band Line-Ups: A Brief Qualitative Look

A crystalline semiconductor has an energy band structure that defines allowed and forbidden energy bands and in the simplest picture, has a conduction band, a valence band and an energy gap separating the two. The exact magnitudes of the energy gap and the conduction and valence band edges (measured with respect to some arbitrary reference level) are determined by the crystal structure and the electrostatic potential of the atomic cores making up the crystal.

Fig. 2.1(a) shows the one-dimensional band diagrams of two semiconductors when they are far apart. Fig. 2.1(b) shows the band diagram when they are brought into contact. To obtain Fig. 2.1(b) from Fig. 2.1(a), we must a priori know the position of the bands in the two semiconductors relative to each other. In other words, we must be able to draw the band diagrams in Fig. 2.1(a) with respect to a common reference energy level. Once this reference level is defined, drawing the band diagram of the heterojunction is straightforward, since we know that the Fermi levels will get aligned. Hence the problem of band line-ups is that of defining a common reference potential when the two semiconductors are widely separated from each other.

Another way of posing the problem is as follows. Since the two energy gaps are unequal, the conduction and valence bands will have discontinuities at the interface. The sum of the two discontinuities will be equal to the difference in the two energy gaps:

$$\Delta E_C + \Delta E_V = E_{G1} - E_{G2} = \Delta E_G$$
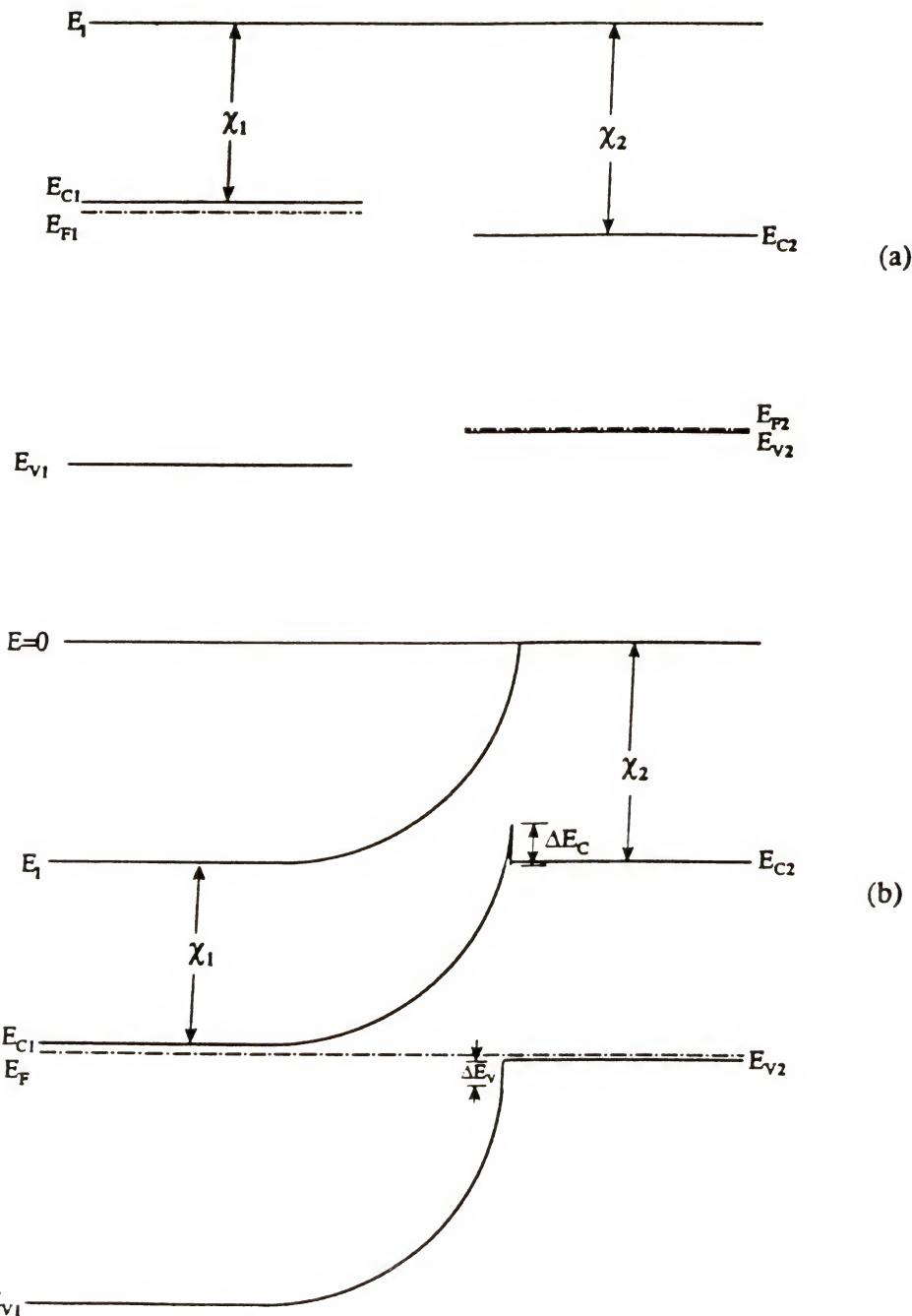
Fig. 2.1 Energy band diagrams of two dissimilar semiconductors when they are (a) far apart, (b) brought into contact.

The way in which ΔE_G gets divided between ΔE_C and ΔE_V is the problem of energy band line-up.

As we mentioned, three approaches exist for determining band line-ups. One is the electron affinity rule, which defines the vacuum level as the reference level for both the semiconductors. The second treats the heterojunction system as a single entity and solves the Schrodinger equation for this system self-consistently to determine ΔE_C or ΔE_V; the third determines ΔE_C or ΔE_V by experimental methods.

2.3 The Electron Affinity Rule

The electron affinity (χ) of a semiconductor is defined as the energy required to take an electron from the bottom of the conduction band to the vacuum level. The electron affinity rule [And62] defines the common reference potential for a heterojunction system as the vacuum level, which is an energy level at which the forces due to the atomic cores become zero. Having defined a reference level, it becomes easy to construct the band diagram of any heterojunction once the electron affinities of the two semiconductors are known. This is illustrated in Fig.s 2.1(a) and 2.1(b). From the way these diagrams are drawn, we can see that

$$\Delta E_C = \chi_2 - \chi_1 \qquad \qquad \text{. . . (2.3-1)}$$

and $\Delta E_V = \Delta E_G - (\chi_2 - \chi_1)$

Thus the problem of band line-ups is solved, quite easily, but at the same time, only theoretically, as we shall soon see. But before that, let us make a digression here.

2.3.1 The Local Vacuum Level: A Digression

Although the electron affinity rule has some fundamental inadequacies, it offers some interesting insights into the potential profiles in and around a semiconductor junction. So let us study it a little more.

Let us perform a thought experiment as illustrated in Fig. 2.2(a), which shows a schematic homojunction between two semiconductors whose band diagram is shown in Fig. 2.2(b). Let us pull an electron from position A in semiconductor 1 to a position B just outside the semiconductor. In doing so, we will impart an energy χ_1 to the electron, by definition of χ_1 (the electron affinity). The electron is now at the so-called local vacuum level (shown as E_1 in Fig. 2.2(b)). Now if we move the electron from B to C, it must travel along the local vacuum level and therefore gain an energy qV_{bi} by the time it reaches C. The fact that the electron changes its energy although it is outside the semiconductor implies the existence of an electric field <u>outside</u> the semiconductor (a fact that is hardly ever mentioned in the literature) in the region where the space charge region exists inside the semiconductor. (The existence of a field should not be surprising; it follows from the requirement of having the tangential component of the electric field continuous at the interface). Thus the local vacuum level effectively represents the variation of the electrostatic potential in, and just outside, the semiconductor. This is a very useful property for heterojunction and variable energy gap semiconductors where the conduction and valence band edges do not always represent the variation
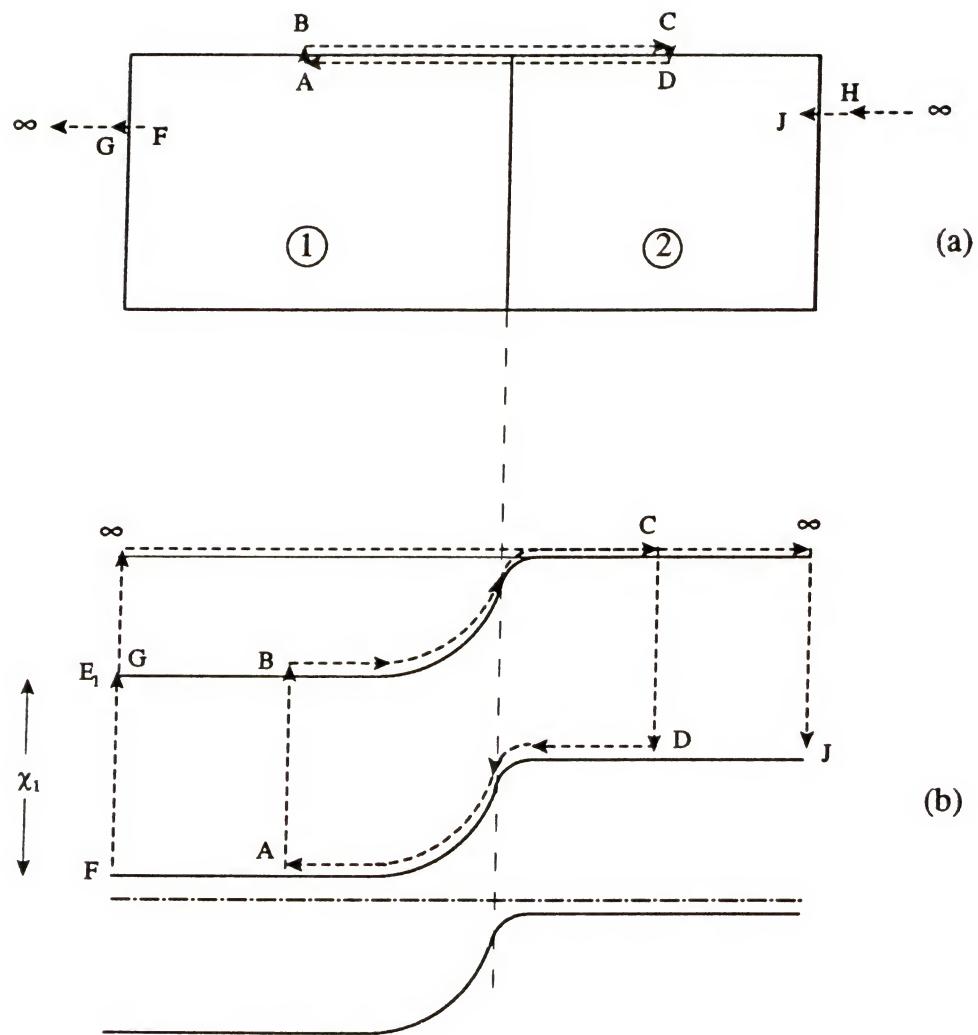
Fig. 2.2 (a) A semiconductor n/p junction and (b) its energy band
diagram. The dashed lines with arrows indicate hypothetical
electron movements.

in the electrostatic potential, so that the local vacuum level provides a convenient means of describing the electrostatic potential. This methodology has been used extensively in device analysis by some authors [Marsh78a, Marsh78b, Lund82].

We must remember though that the local vacuum level is only a theoretical quantity and is never measured in practice. In fact, it is possible that, in any actual semiconductor structure, an electric field in the region surrounding the semiconductor may cause the ionic species in the air to drift and settle on the semiconductor surface, thus forming a dipole layer until the field outside becomes zero. But from the viewpoint of a theoretical device physicist, whether this happens or not is irrelevant.

As a further digression, let us do another thought experiment. Let us pull an electron from position F (Fig. 2.2(a)) out of the semiconductor to G, then to infinity, and then bring it back on the p-side to H. As shown in Fig. 2.2(b), in going from F to G, the electron will gain an energy χ_1. But then, if we define the zero of the energy as the E_1 on the p-side (Fig.2.2(b), the electron will gain an additional energy $(=qV_{bi})$ as it goes from G to infinity, although it is outside the semiconductor during the whole transport! This implies that there must be an attractive force on the electron <u>outside</u> the semiconductor. We thus conclude that whenever a junction is formed that creates a space charge region at the metallurgical junction, a charge would simultaneously be created on the surfaces that would create a field outside the semiconductor. The polarity of this charge will be the same

as the polarity of the charge on that side of the metallurgical junction.

2.3.2 More about the Electron Affinity Rule

The definition of electron affinity, namely, the energy required to take an electron from the bottom of the conduction band to the vacuum level, has been given two different interpretations.

One approach interprets the definition as the energy required to physically remove an electron from the semiconductor. This process involves overcoming two potentials. First, the binding energy of the electron necessary to overcome the correlation and exchange potentials; and second, the energy to overcome the surface dipole layer potential [Bar36]. The dipole layer at the surface arises due to the sudden discontinuity in the lattice, which causes the electron clouds around the surface atoms to be asymmetric. The calculation of this potential requires solving the Schrodinger equation self-consistently [Ben67, Wig35, Bar36, Flo77].

The second approach interprets the definition simply as the binding energy of the crystal, and thus does not involve any surface considerations. We discuss both these approaches below.

According to the first interpretation, the electron affinity (χ) can be written as

$$\chi = B.E. + D^S$$

where

B.E. = binding energy of the crystal

and

D^S = surface dipole potential

Bennett and Duke [Ben67], Kroemer [Kro75, Kor83a] and others have expressed reservations about using this interpretation to determine energy band line-ups. The following is the gist of their argument.

When a heterojunction is formed from two dissimilar semiconductors, the atoms at the interface form bonds with each other, and thus most of the dangling bonds that exist at a free semiconductor surface are satisfied. The properties of this interface are, therefore, chemically (and therefore electronically) very different from those of a free semiconductor surface. But the electron affinity rule asserts that (Eq. (2.3-1)),

$$\Delta E_C = \chi_2 - \chi_1$$
$$= (B.E.)_2 + D^S{}_2 - (B.E.)_1 - D^S{}_1 \tag{2.3-2}$$

Such a formulation implies that the properties of a heterojunction interface can be obtained simply by superimposing the properties of the constituent semiconductor-vacuum surfaces. It thus neglects the change in the properties due to the chemical bonding between atoms of the two semiconductors.

From another point of view, due to the bonding between atoms, the problem of heterojunction energy band line-ups must be solved within the bulk of the semiconductors, and the problem has no relation to the surface properties of the constituent semiconductors.

Shay et al. [Sha76] pointed out that in practice, the definition of Eqn. (2.3-2) may still be a good approximation because they found the

surface dipole contributions to the electron affinity to be small. But other researchers have disagreed with this claim [Bau83, Kro83a].

According to the second interpretation of the definition of electron affinity (χ), it is simply written as

$$\chi = B.E.$$

Thus the discontinuity in and conduction band at a heterojunction will be

$$\Delta E_C = \chi_1 - \chi_2 = (B.E.)_1 - (B.E.)_2$$

Such a definition is theoretically rigorous [Sah91], since the binding energy is a bulk property of a crystal, and thus is independent of the surface properties. The question that must be answered with regard to this definition is whether it is commensurate with experimental methods or not; since experimental methods of determining the electron affinity would presumably involve the physical removal of electrons from the semiconductor. The physical removal, in turn, would involve overcoming the surface potentials.

Research continues to be done on the determination of band line-ups using self-consistent solutions of Schroedinger's equation. The electron affinity rule continues to be used by many device physicists such as Marshak and van Vliet [Marsh84], Lundstrom [Lun84], Sah [Sah91] etc.

We mention another approach suggested for determining band lineups by Adams and Nuusbaum [Ada79]. They proposed that when two dissimilar materials are brought together, the intrinsic Fermi levels (E_i) will line up; and they had a mathematical proof for their

assertion. The proposal caused an interesting debate in the literature from various device physicists [Kro83b, Nus84, Roo85], until finally it was realized [Lee85], by a graduate student, that Adams and Nussbaum had made a sign error in their derivation, and thus their assertion was invalid. The proposal, therefore, was finally discarded by Nussbaum himself [Nus85].

For the purposes of the present work, we shall assume that the band line-up of the heterojunction under consideration is known apriori, and we shall not worry about how it was arrived at.

CHAPTER 3
CURRENT TRANSPORT MECHANISMS IN HETEROJUNCTIONS

3.1 Introduction

There are many different physical mechanisms of electron and hole transport found in heterojunctions depending on the characteristics of the interface. These mechanisms are drift and diffusion, thermionic emission, tunneling, interface recombination, space charge region recombination, and any combination of the above.

In this chapter, we will review the literature on aspects of this subject in Sec. 3.2; and then take a brief tutorial look at the current flow mechanisms of thermionic emission in Sec. 3.3 and tunneling in Sec. 3.4. As explained at the end of chapter 1, this is a review chapter. It is included here to (i) get an idea of the evolution of the understanding of current transport mechanisms in heterojunctions (Sec. 3.2) and (ii) to present this author's understanding of two important transport mechanisms, which would be useful for better appreciating the models presented in later chapters. It is also hoped that the section on thermionic emission (Sec. 3.3) will provide one more point of view on the literature on this subject, which this author has found to be very cryptic and unsatisfying in providing physical insight.

3.2 Literature Review

Although altogether different in many respects, the metal-semiconductor (M/S) junction and the semiconductor-semiconductor heterojunction do share some common features; and since much more work

has been done on the M/S junction than on the semiconductor
heterojunction, a student of the latter will learn much by referring to
the literature on the former. The reader is therefore referred to the
books by Henisch [Hen84], Rhoderick [Rho80] and Sze [Sze81] for a
comprehensive description of (and hundreds of references on) current
flow mechanisms in M/S contacts. Below we review the literature on
current flow mechanisms in heterojunctions.

The first paper (known to this author) written on this subject
was in 1962, by Anderson [And62], who reported measurements on n-Ge/P-
GaAs heterojunctions.[1] He proposed that the current flow across the
heterojunction could be described by the conventional drift-diffusion
mechanism as in a homojunction. The analytical model in the paper was
rather brief and sketchy. (It was in this paper that the electron
affinity rule was first used to construct the band diagram of a
heterojunction, a diagram that was to set the standard for band line-ups
for the next twenty or so years).

In 1964, Perlman and Feucht [Per64] proposed another model to
better fit the data on n-Ge/P-GaAs junctions similar to those fabricated
by Anderson. The model proposed a thermionic emission flow of carriers
over the abrupt heterojunction barrier, as in a Schottky barrier,
flowing by diffusion in the quasi-neutral regions (Fig.3.1(a)). The
balance between the two currents would determine the I-V
characteristics. The authors derived analytical expressions for the

[1]We shall use the convention of indicating a junction by a "/" ; and in
a heterojunction, the dopant type of the wider gap material will be
denoted by a capital letter. Thus n-Ge/P-GaAs indicates a
heterojunction between n-type germanium and p-type gallium arsenide,
with gallium arsenide having the larger of the two energy gaps.

thermionic emission flux and the diffusion flux and then set criteria for which current will dominate under what conditions. (A similar model for the M/S junction was derived two years later by Crowell and Sze [Crowe66]).

At about the same time, Rediker et al. [Red64] reported measurements on Ge/GaAs and GaAs/GaSb heterojunctions, and showed that the diffusion theory of Anderson did not agree with their experimental results. These experimental results indicated that the I-V characteristics were of the form $I = I_O \exp(AV)$ where A was a constant relatively independent of temperature. A diffusion dominated current would have required an exponent of the form (qV/nkT) which is strongly temperature dependent. The authors therefore concluded that the dominant mechanism must be tunneling through the spike (Fig. 3.1(b)), because similar I-V characteristics were observed for m/i/m structures [Adv62]). They also stated a very approximate expression for the tunneling probability across the barrier.

But then in 1966, Riben and Feucht [Rib66a, Rib66b] did numerical evaluation of the tunneling integrals involved and found that the effective mass of electrons required to fit the experimental data would have to be unrealistically small; thus tunneling through the spike had to be discarded as a plausible current flow mechanism (in n-Ge/P-GaAs junctions). Taking a clue then from tunnel diode theory [Chy61, Sah61a], they proposed a tunneling-recombination model where carriers were assumed to get trapped into energy gap states at the interface and then
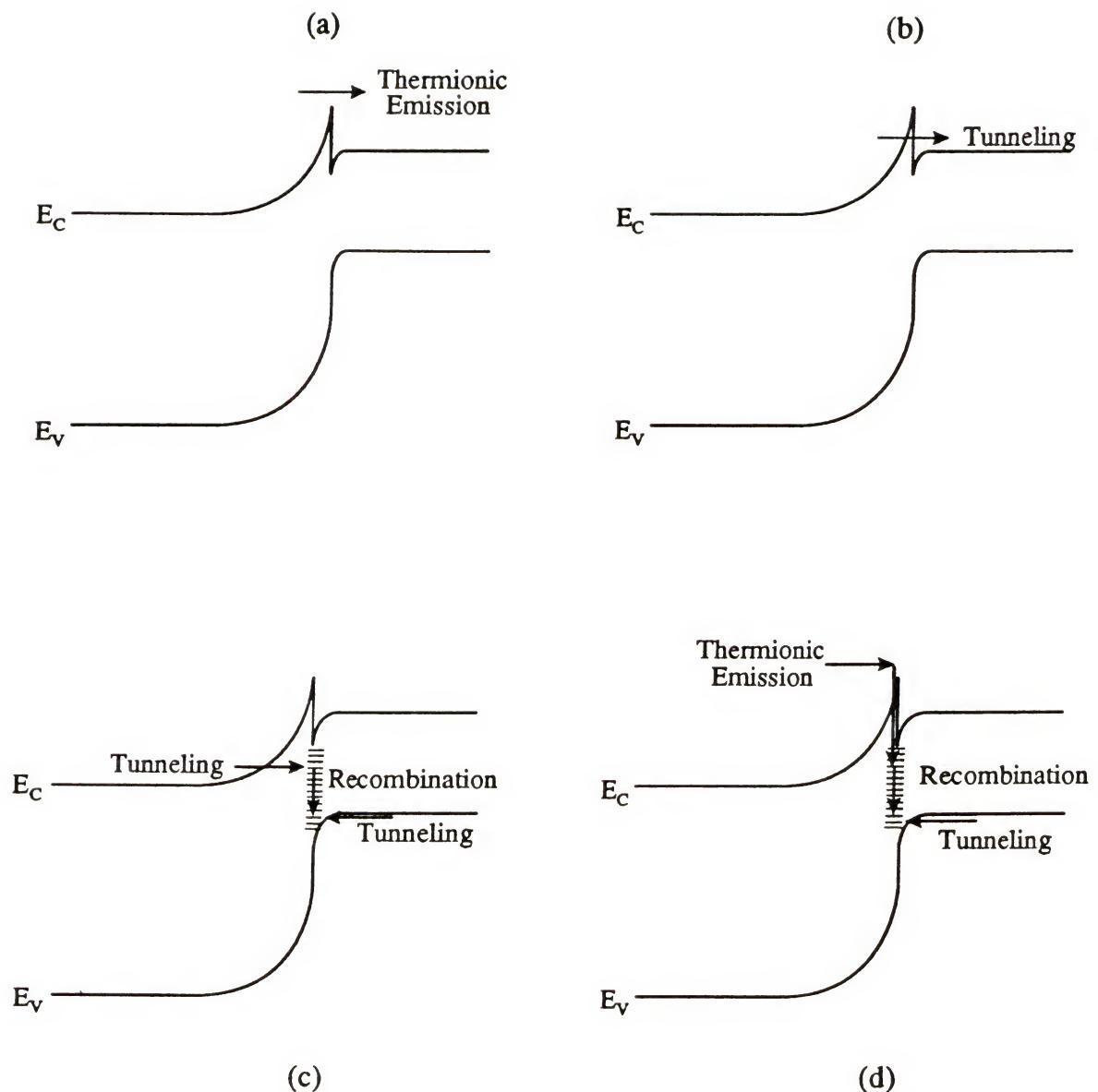
Fig. 3.1 Current flow mechanisms across a heterojunction.

tunnel into the other band (Fig. 3.1(c)). They also suggested that this process may consist of multiple tunneling and trapping steps.

Concomitantly, Donnelly and Milnes reported [Don66] that the forward I-V characteristics of p-Ge/N-GaAs junctions exhibited two regimes of operation: 1. A low bias regime in which $I \propto \exp(qV/nkT)$; and 2. a high bias regime for which $I \propto \exp(AV)$, where A was a constant practically independent of temperature. The low bias regime therefore indicated a diffusion dominated current, as first proposed by Anderson [And62], while the high bias regime indicated a tunneling dominated current. To explain these characteristics, they proposed a three step carrier flow process consisting of the following components (Fig. 3.1(d)):

1. Tunneling of majority carriers of the narrow gap material (holes in Ge in their case) through the barrier into the interface states;
2. recombination of the above trapped carriers (holes) with the majority carriers (electrons) of the wide gap material (GaAs); and
3. majority carrier diffusion-drift in the space-charge region of the wide gap material.

These three carrier fluxes will flow in series, and the smallest of the three will determine the magnitude of the current that will ultimately flow. The authors derived zeroth order expressions for the three carrier fluxes which must be equated and solved for (in steady state), to find the final I-V characteristics. But the equations were too complex to be solved in a closed form, and therefore they discussed only some special case solutions.

Halil and Kao wrote a paper [Hal75] in 1975 on heterojunction tunnel diodes in which they formulated a tunneling-recombination model similar to that of Riben and Feucht [Rib66a]. Unlike the latter, their mathematical treatment was rigorous; but the integrals became so complex that they ended up writing their final I-V relationships in terms of some arbitrary fitting parameters. Thus they did not really utilize the rigorous formulation that they started off with. An interesting aspect of their formulation was that they treated the band state-to-bound state tunneling process in a fashion identical mathematically to the carrier trapping process. Thus they wrote the conduction band to trap state tunneling flux as

$$dU_T = [\ Z_{ct}(E)g_c(E)f(E)N_t(E)(1-f_t) - Z_{tc}(E)g_c(E)(1-f)N_tf_t \ ] \ dE$$

where Z_{ct} and Z_{tc} are the tunneling probabilities from the conduction band to the trap and vice versa. This paper provides a good starting point for building a model for the tunneling-recombination current, but by itself, it is far from complete.

Smith [Smi77] formulated a model which extended the famous Sah-Noyce-Shockley theory [SahC57] for homojunctions to heterojunctions by including interface recombination. The model was theoretical/numerical, and did not provide experimental verification. Thus it was not clear considering the many current mechanisms possible in a heterojunction, as to the conditions under which this model would be applicable.

The above mentioned papers constitute all the work that has been done on modeling current flow mechanisms in heterojunctions. Except the last two, all the papers proposed models to explain experimental

results. Derivation of rigorous analytical models was not emphasized, and therefore the models described were, at best, semi-empirical. Experimental investigations done today still refer back to these early papers to figure out the dominant current flow mechanisms. Thus the status today is that all the mechanisms that may contribute to current flow at a heterojunction are known, along with a general understanding of how each of these varies with various material properties and operating conditions. Thus one has the ability to explain experimental results of any heterojunction by comparing them with the known characteristics of different current flow mechanisms. But lacking are rigorous analytical models for the various current flow mechanisms. Still more importantly, a unifying theory encompassing all the different mechanisms, that will enable one to predict the I-V characteristics of any heterojunction, is needed. It is the hope of this author that such a theory will be developed in the future.

Papers on quantum mechanical studies of current transport across heterojunctions were found [WuC79, Osb79], but not studied in detail. There are also papers on isotype heterojunctions (that is N/n or P/p junctions) [Chan80, LeeS81, Mae86, Tho87]. Apart from papers on analytical modeling, there have of course been many papers on the experimental characterization of various heterojunctions (e.g. [Dun73, Gas73, HallW73, Ota73, Boc74, Mart76, Osi76, Lup77, Nel78a, Nel78b, Gri79, Chan79, Vla80, Gor80, Shi83, Cha85a, Fis86, Ima87, Lyo89a, Lyo89b], and many others).

3.3 Thermionic Emission

We do a short tutorial study of the thermionic current flow across a heterojunction barrier in this section. For concreteness, we shall study the transport of electrons. The study of holes will be analogous.

Consider a bulk semiconductor in steady-state, in which a gradient of electron density exists along the x-axis, as shown in Fig. 3.2. Such a gradient will cause a net electron flow in the positive x-direction, due to the process of diffusion. Diffusion is the result of the random thermal motion of electrons. For example, consider a plane S perpendicular to the x-axis, as indicated in Fig. 3.2. Electrons will be constantly crossing this surface from left to right and right to left due to their thermal motion. But since there are more electrons to the left of the surface, on the average more electrons will cross over to the right than to the left, thus causing a net diffusive flow.

Let us derive here the expression for the diffusion flux, F_D. Let the surface S be located at $x=x_0$. On the average, we can assume that all electrons have a component of velocity v_x along the x-axis. Now consider two differential volumes a small distance Δx on either side of S. Let Δx be less than the mean free path, λ. The electron densities within these volumes will be

$$n(x_0+\Delta x) = n(x_0) - \Delta x \, dn/dx \qquad (3.3\text{-}1a)$$

and
$$n(x_0-\Delta x) = n(x_0) + \Delta x \, dn/dx \qquad (3.3\text{-}1b)$$

where $dn/dx < 0$. Thus the net flux, dF_D, crossing the plane S, due to the electrons within these two volumes will be
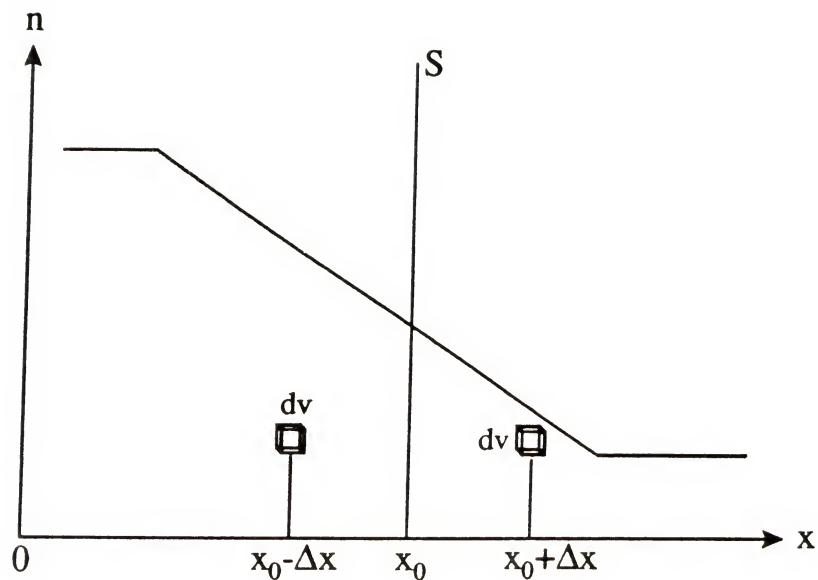
Fig. 3.2 Plot of an electron density gradient in a bulk semiconductor. S is an imaginary surface perpendicular to the plane of the gradient.

$$dF_D = v_x \ n(x_0 - \Delta x) - v_x \ n(x_0 + \Delta x) \qquad (3.3\text{-}2)$$

$$= -2 \ v_x \ \Delta x \ dn/dx \qquad (3.3\text{-}3)$$

The net electron flux, F_D, crossing S per second can be found by integrating the flux dF and then calculating the average distance from which electrons will be able to reach S within one second [Ken38]. In the calculation, one assumes that a collision totally randomizes the velocity of the electron; and thus integration is done over only those electrons which have undergone their last collision before crossing S. The mathematics of the calculation is complicated, and is not repeated here [Ken38]. The final result is found to be

$$F_D = -v_{th} \ \lambda/3 \ dn/dx \qquad (3.3\text{-}4)$$

where v_{th} is the electron thermal velocity, and λ is the mean free path. The quantity $v_{th}\lambda/3$ is the diffusivity, D. Diffusion therefore is seen to be a process intimately tied to the mean free path.

Now let us consider a metal/semiconductor heterojunction as shown in Fig. 3.3(a). There exists a sharp gradient in the electron density at the interface (Fig.3.3(b)). Note two points about this situation: (i) The length over which the gradient exists is much smaller than the mean free path of electrons; and (ii) an energy barrier is necessary to maintain such a sharp gradient without causing a large current to flow.

Will electrons "diffuse" across the interface in such a situation? No, they will not. Why not? For the following reasons.
1. The presence of the energy barrier prevents most of the electrons on either side from participating in the flow across the interface. Thus

only those electrons that have kinetic energies higher than the barrier height can cross over.

2. Since the thickness of the interface region ($\simeq$ 20 °A) is small compared to all other dimensions of interest, the interface can be assumed to be abrupt for all practical purposes. Thus the gradient in n will be zero everywhere except at the interface, where it will be very large. Thus the electron flux, F_N, across the interface can be written, as in Eqn. (3.3-2), as

$$F_N = v_x \, n(x=0^-) \, \exp(-\Phi_{MS}/kT) \, - \, v_x \, n(x=0^+) \qquad (3.3-5)$$

where Φ_{MS} -- the barrier height -- is defined in Fig. 3.3(a). But the flux cannot be written further in terms of a gradient dn/dx, as we did in Eqn. (3.3-2), since dn/dx is now ill-defined. Hence electron flow across an interface where an abrupt barrier exists must be described by an equation such as (3.3-5), but not by the equation that describes diffusion. Diffusion is a meaningless term when the density gradient exists in a region smaller than the mean free path.

The process described by Eqn. (3.3-5) is thermionic emission. Classically, the term thermionic emission was used to describe the emission of electrons from heated metal surfaces. In that phenomenon, the semiconductor of Fig. 3.3(a) will be replaced by air.

Another analogy to thermionic emission is the phenomenon of effusion that occurs in gases: Consider a chamber divided into two regions by a partition which has a hole in it. Let the size of the hole be small compared to the mean free path of the gas molecules. Let one
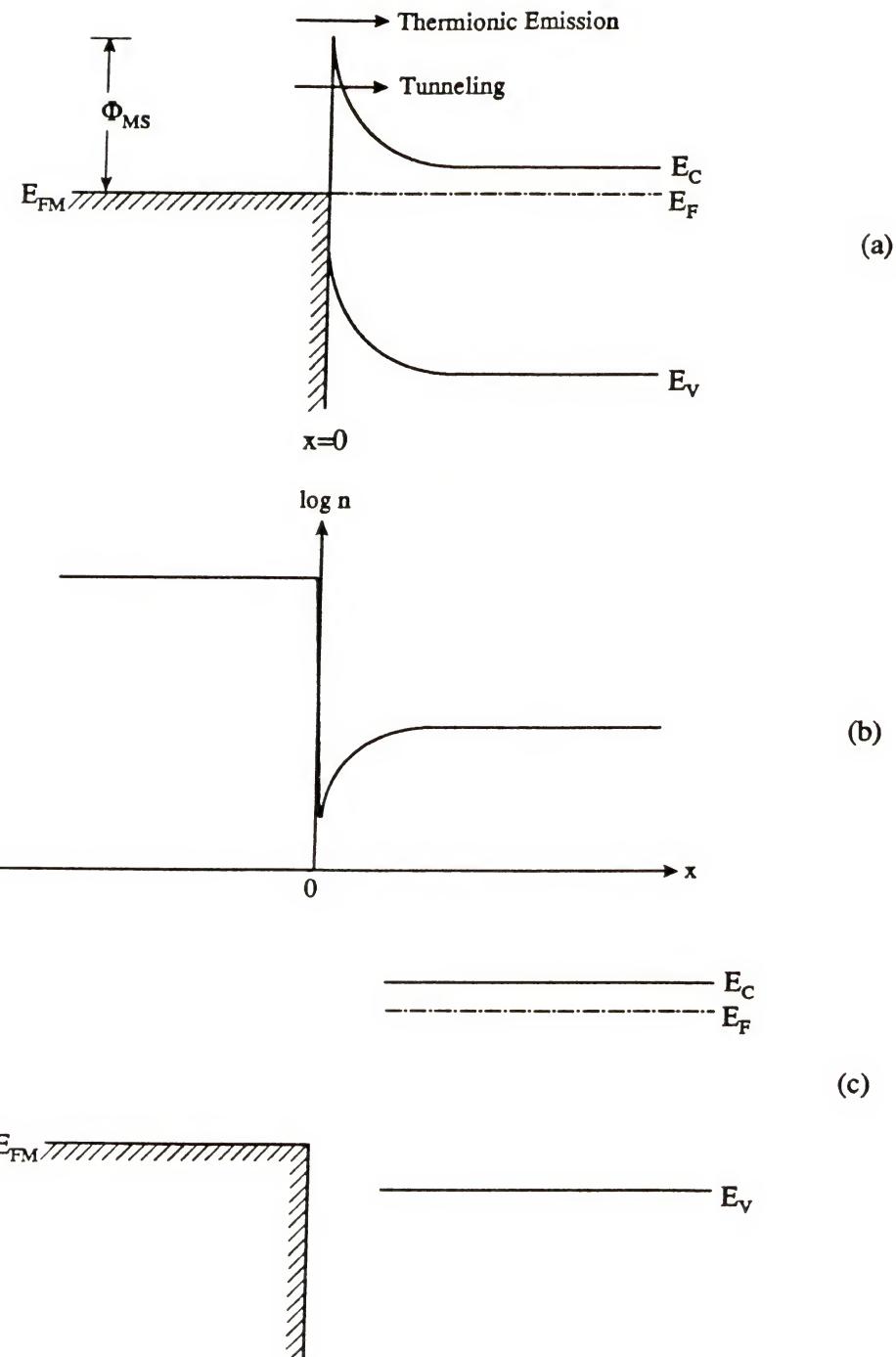
Fig. 3.3 A metal/semiconductor junction: (a) energy band diagram at equilibrium, also showing the electron flow mechanisms; (b) electron density as a function of position; (c) energy band diagram when the two materials are far apart.

side of the partition be filled with an ideal gas. The other side may or may not have a gas in it. Only those gas molecules that happen to approach the hole will pass through. The rate of flow will be independent of the gas pressure on the other side of the partition (unlike in diffusion), since the size of the hole is so small that the gas remains "unaware" of the pressure on the other side. The large partition is equivalent to the energy barrier in case of thermionic emission, both of which prevent a majority of the carriers from participating in the flow. The size of the hole corresponds to the inverse of the barrier height: The smaller the hole (the larger the barrier height), the lesser will be the number of carriers going through.

A brief but interesting description of thermionic emission can also be found in Crowell and Hafizi's paper [414]. Some other informative references are [Rho80, Hen84, Crowe66, Per64, WuC82].

Digression: We note another interesting aspect of Fig. 3.3. Although the electron density in the metal is much higher than that in the semiconductor, electrons are transferred from the semiconductor to the metal, when the two are brought in contact. Such flow is opposite in direction than the direction one might expect a diffusive flow to occur. The reason here is that electrons diffuse not necessarily from a region of higher concentration to a region of lower concentration, but rather from a region of higher chemical potential (that is the Fermi potential in metals and semiconductors) to regions of lower chemical potential. This can be proved from thermodynamic arguments: First, the chemical (or

Fermi) potential, by definition, is the change in the Gibbs free energy of a system (metal of semiconductor) when an electron is added to it. Thus mathematically,

$$E_F \triangleq dG/dn$$

where E_F = Fermi potential, G = Gibbs free energy, and n = electron density. Second, the second law of thermodynamics states that in any irreversible process, the Gibbs free energy of the system will decrease. It is easily deduced from these two statements that when two materials having different Fermi levels when far apart (see for example Fig. 3.3(c)) are brought into contact, electrons will flow from the material having a higher Fermi level to that having a lower Fermi level, since it is flow in this direction that would cause the Gibbs free energy of the system to decrease. The notion, therefore, that diffusive flow of electrons is in the direction of decreasing electron density, is correct only in a material of homogeneous composition.

3.4 Tunneling

If a spike exists in the conduction band at the heterojunction interface, then apart from electrons being emitted thermionically over the energy barrier, there are electrons tunneling through the barrier also. Das and Lundstrom [Das88] showed, using numerical calculations, that such tunneling can form a significant fraction of the total current flowing across the heterojunction interface. Thus a model for current flow across the heterojunction interface must include the tunneling current.

Historically, the tunneling of electrons through energy barriers at metal or semiconductor surfaces has also been called "field emission" (to be contrasted with thermionic emission). At a metal surface, for example, electrons can escape out from the metal due to their thermal energy. This is thermionic emission. But if a large electric field is present at the surface, then the energy barrier at the surface will get lowered, and electrons can then tunnel through the barrier; hence the name field emission. The mechanisms are as illustrated in Fig. 3.3(a).

Analytical modeling of the field emission phenomenon was first done by Fowler and Nordheim [Fow28] and Nordheim [Nor28]. Since then many more formulations have been proposed, both for metals and semiconductors [Str55, Mur56, Str62, Pad66, Con66, Chr66, Chr67, Crow69, Rid70, Chang70]. Of these, the most comprehensive theory was developed by Christov, both for metals [Chr66] and for semiconductors [Chr67]. The model that he developed included a wide range of temperatures (0-3500 K) and electric fields (0-10^8 V/cm). Also, it did not just model the flow of carriers due to tunneling, but rather the total carrier flow across the energy barrier. Thus it accounted for tunneling, thermionic emission, and the quantum mechanical reflection above the energy barrier. The mathematics of the model was extremely involved, and several different expressions were obtained for different energy ranges and conditions of operation. We shall not study the modeling of tunneling here. We just mention one aspect of the methodology that is relevant to this work. The total electron current density flowing across the heterojunction barrier is written as [Chr67]

$$J_N = q \ 2/h^3 \ \iiint f(E) \ W(E) \ \partial E/\partial p_x \ dp_x \ dp_y \ dp_z \qquad (3.4\text{-}1)$$

where

h = Planck's constant

E = kinetic energy of electrons

f(E) = probability of occupation

W(E) = probability of transmission at energy E

p_x, p_y, p_z = components of electron momentum

Note also that $\partial E/\partial p_x$ is the electron velocity in the x-direction, and $(2/h^3)$ is the density of electron states in the conduction band expressed in momentum space. Noting that the electron density n can be written as

$$n = 2/h^3 \ \iiint f(E) \ dp_x \ dp_y \ dp_z \qquad (3.4\text{-}2)$$

Eqn. (3.4-1) can be written as

$$J_N = q \ n \ S_N \qquad (3.4\text{-}3)$$

where

$$S_N = \frac{2/h^3 \ \iiint f(E) \ W(E) \ \partial E/\partial p_x \ dp_x \ dp_y \ dp_z}{2/h^3 \ \iiint f(E) \ dp_x \ dp_y \ dp_z} \qquad (3.4\text{-}4)$$

S_N is an effective velocity that describes the transport across the heterojunction barrier; and as expressed in Eqn. (3.4-4), it includes both thermionic emission and tunneling fluxes. We will be using this concept in chapters 4 and 5.

This concludes our review of current flow mechanisms in a heterojunction structure.

CHAPTER 4
REVIEW OF HETEROJUNCTION BIPOLAR TRANSISTOR (HBT) LITERATURE

In this chapter we review the literature on (i) analytical models for the HBT (Secs. 4.1 through 4.5); (ii) hot electron transport through the base of the HBT (sec. 4.6); (iii) velocity overshoot and its exploitation within the base-collector space charge region of the HBT Sec. 4.7); and (iv) HBT fabrication technology (Sec.4.8).

4.1 Review of HBT Models

All the models proposed for HBTs are extensions of the Ebers-Moll [Ebe59] or the Gummel-Poon [Gum70b] models. The one thing that distinguishes an HBT model from a homojunction bipolar transistor model is that the former must, in some way, account for the effects of the heterojunction on the transistor characteristics. There are two ways that the heterojunction affects device characteristics: (i) it affects the current flow across the junction, as we discussed in chapter 3; and (ii) it affects the injection efficiency of the junction that is a heterojunction. Of these, (ii) is quite straightforward, almost trivial, to account for [Kro57]. Thus all HBT models concentrate primarily on modeling the current flow across the heterojunction and incorporating that current model into the model for the transistor.

There have been four major models proposed for the dc characteristics of the HBT [Marty79, Gri84, Lun86, Ryu90], all of them for GaAlAs/GaAs HBTs. The first [Marty79], proposed in 1979, subsequently extended to double heterojunction bipolar transistors

(DHBTs) [Bai85] and compared with experimental results [Tas86], was a Gummel-Poon like charge-control model with the current flow across the heterojunction assumed to be due to the conventional drift-diffusion mechanisms. The second [Gri84], proposed in 1984, was a thermionic emission-diffusion-tunneling model, extended to DHBTs in 1986 [LeeS86]. This was a comprehensive formulation for current flow across a heterojunction in which interface recombination can be neglected. The paper also extended the model to graded energy gap emitter transistors. The formulation was as that of the Ebers-Moll model. The third model [Lun86], proposed in 1986, was also an Ebers-Moll like model; it left the current flow mechanism at the heterojunction unspecified in terms of a generalized interface transport velocity. The fourth [Ryu90], proposed in 1990 was (like the first [Marty79]) a Gummel-Poon like model, and utilized the generalized interface transport velocity concept of [Lun86]. It was thus a more general model than [Marty79]. It also took into account bulk and surface recombination within the heterojunction space charge region, and high current effects in the collector. We review these four models in the next four sections.

There are many aspects in which homojunction and heterojunction transistors exhibit identical behavior; thus familiarity with the homojunction bipolar transistor (BJT) literature is an essential pre-requisite for an HBT researcher. Hence we list here the literature on BJTs that this author referred to. The list is, of course, not comprehensive. Papers on charge-control models [Gra71, Win73, Vau77, Graa85, Sch85, Ben86, McD88, Rei88], on effects at high current

densities [Vlie74, Jeo87, Stu87, Rei87, Yuan88b], on high frequency characteristics [Kwe76, Rou86, Nan88, Bie88], on transit or delay times [Kau78, Sin79, Bie86, Mey87, Neg88, Jos88], and other papers addressing specific issues [Gia74, Daw75, Sir76, How78, WuC80, Fos81, Moh87, Sei87, Heb88, Hur88, Yuan88a] were encountered. Two books [Eng86, Gra90] provide good reviews of analytical models for transistors.

We also mention here papers that have reported numerical studies on heterojunction bipolar transistors: [Asbe82, Kur84, Yok84, Yos85, Kat87, Ank86, Kat89a, Kat89b, Mey89, Hor90].

Throughout this chapter, we will use the nomenclature defined in the band diagram of Fig. 4.1. The figure shows an abrupt N-AlGaAs/p-GaAs/N-AlGaAs HBT. The various symbols have the following meanings. ΔE_C and ΔE_V are the discontinuities in the conduction and valence bands respectively (for a discussion of energy band alignments at a heterojunction, see chapter 2). V_{JP} and V_{JN} are the electrostatic potential drops across the space charge layer, on the p and N sides, respectively. X_{NE} and X_{pE} are the space charge layer thicknesses on the N and the p sides respectively, of the emitter-base junction; X_{NC} and X_{pC} are similar quantities at the collector-base junction. The zero of the x-axis is defined to be at the metallurgical interface of the emitter-base heterojunction. W_E, W_B and W_C are the metallurgical thicknesses of the emitter, base and collector layers. E_{FN} and E_{FP} are the electron and hole quasi-Fermi levels.
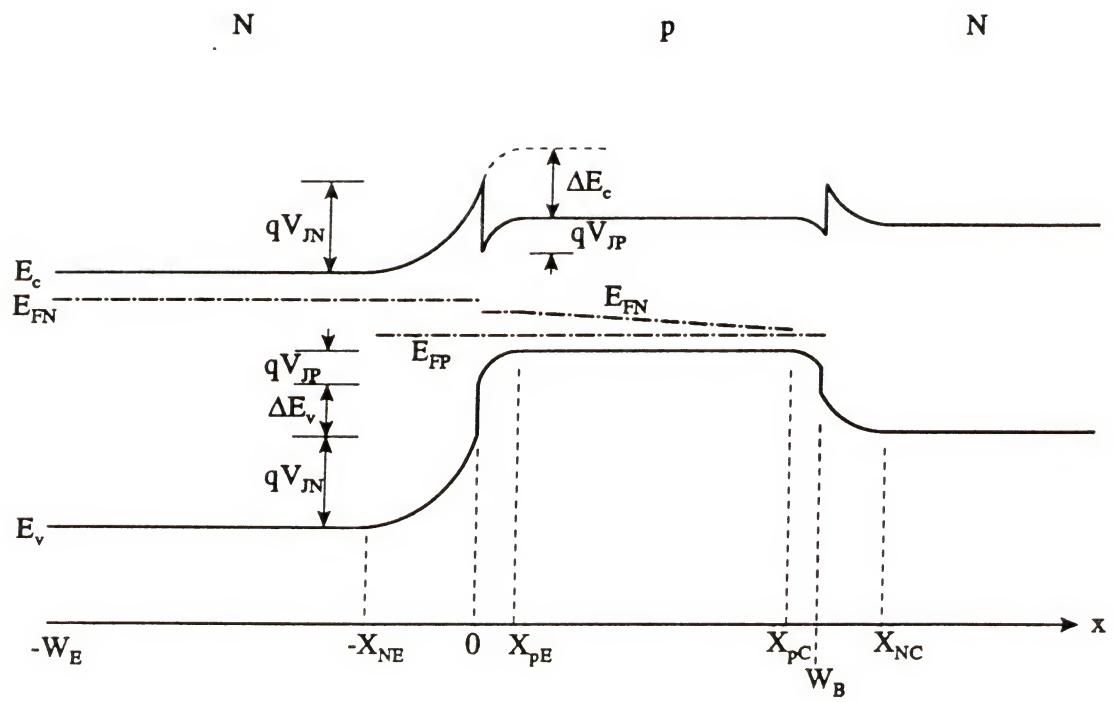
Fig. 4.1 Energy band diagram of an abrupt double heterojunction bipolar transistor.

4.2 The Charge Control Model

This is a model for an abrupt N-AlGaAs/p-GaAs/n-GaAs

heterojunction transistor [Marty79], and it concentrates mainly on

modeling the effect of the spike in the conduction band on the I_C-V_{BE}

characteristics. The collector current density (J_C) in the model is

derived to be

$$J_C \simeq -\frac{\left(\dfrac{pn}{n_{iB}^2}\right)_{-X_{NE}} - 1}{\displaystyle\int_{-X_{NE}}^{W_B} \frac{p}{qD_n n_{iB}^2}\, dx} \qquad (4.2\text{-}1)$$

where n_{iB} is the intrinsic carrier density in the base region, and the

other symbols have their usual meanings. The methodology used to derive

the model is similar to that of Gummel's charge-control model [Gum70a].

The three major assumptions made in the model are: (i) The base-

collector junction remains at equilibrium, that is, the base-collector

bias is always zero; (ii) there is no recombination anywhere in the

device; and (iii) the current flow across the heterojunction is by

drift-diffusion alone.

Assumption (i) restricts the use of the model to the active

regime of operation. Assumption (ii) causes two major restrictions: (a)

Neglecting recombination in the quasi-neutral base implies that the

current gain β must be dominated by the emitter injection efficiency.

This is in general true for abrupt junction HBTs, but still it restricts

the utility of the model considerably. (b) Neglecting the emitter-base

space charge layer recombination prevents the use of the model for low
current levels when such recombination is dominant. But assumption (iii)
is perhaps the most drastic one. In an abrupt heterojunction, if the
spike in the conduction band is larger than a few kT, then current flow
will be by thermionic emission, as we discussed in Sec. 3.3. Thus using
drift-diffusion to describe this current flow is simply incorrect.
Further, the experimental results reported in the paper are also not
explicable by the proposed model, as the authors themselves admit. In
particular, the temperature dependence of the ideality factor n in the
equation

$$I_C = I_{C0} \exp (qV_{BE} / nkT)$$

(n decreases with increasing temperature), is unexplained by the model.
We believe that the experimental results would be easily explained if
the tunneling-recombination model proposed by Riben and Feucht [Rib66a,
Rib66b] is used.

Thus in conclusion, it seems that this model does not have much
merit.

4.3 Thermionic Emission-Diffusion-Tunneling Model

From the point of view of current flow across the heterojunction
interface, this is the most comprehensive analytical model proposed to
date [Gri84]. The current is assumed to flow by thermionic emission-
tunneling at the interface and as drift-diffusion in the bulk of the
semiconductor. The collector current density is derived as

$$J_{NC} = \frac{\dfrac{-qD_n \, n_{0B}}{L_n \sinh(W_B/L_n)} \left\{ [\exp(qV_{BE}/kT)-1] - [\exp(qV_{BC}/kT)-1]\{\cosh(W_B/L_n) + \atop + \eta_n \, \sinh^2(W_B/L_n) \, \exp(-\Delta E_n/kT) \} \right\}}{1 + \dfrac{(D_n/L_n) \, \coth(W_B/L_n)}{v_x \gamma_n \, \exp(\Delta E_n/kT)}}$$

$$\dots (4.3\text{-}1)$$

where

$$\Delta E_n = qV_{JP} - \Delta E_C,$$

L_n is the diffusion length, v_x is the mean thermal velocity of electrons ($v_x = (kT/2\pi m_n^*)^{1/2}$),

$$\eta_n = \frac{4D_n}{v_n \, L_n \, \sinh(W_B/L_n)} \qquad (4.3\text{-}2)$$

and γ_n is the integrated probability of tunneling through the conduction band spike. If the term involving η_n in the numerator is negligible, then the numerator will be exactly the expression for the collector current of a homojunction transistor. The denominator represents the effect of the interface carrier flow on the current. Let

$$D \overset{\Delta}{=} 1 + \frac{(D_n/L_n) \, \coth(W_B/L_n)}{v_x \gamma_n \, \exp(\Delta E_n/kT)} \overset{\Delta}{=} 1 + \frac{v_d}{v_{tf}}$$

where, v_d = effective diffusion velocity in the base and v_{tf} = effective thermionic field emission velocity across the emitter-base interface.

Thus if $v_{tf} \gg v_d$, then $D \approx 1$ and the junction will behave as a homojunction. On the other hand, if $v_{tf} \ll v_d$, then the junction will behave as a metal-semiconductor junction having a large barrier height [Crowe66]. Now the magnitude of v_d is,

$$v_d = (D_n/L_n) \, \coth(W_B/L_n)$$

Thus for a transistor with $W_B \ll L_n$,

$$coth(W_B/L_n) \simeq L_n/W_B$$

so that

$$v_d \simeq D_n/W_B$$

which indicates that the effective diffusion velocity depends inversely on the base thickness. On the other hand, for a thick base diode with $W_B \gg L_n$,

$$coth(W_B/L_n) \simeq 1$$

so that

$$v_d \simeq D_n/L_n$$

Thus it is possible to imagine that the same heterojunction will exhibit homojunction-like characteristics as a thick base diode, but show metal/semiconductor junction-like characteristics for a narrow base transistor.

The magnitude of v_{tf} is

$$v_{tf} = v_x \gamma_n \exp(\Delta E_n/kT)$$

For an N-AlGaAs/p-GaAs junction, ΔE_n is a negative quantity. For a heavily doped base

$$\Delta E_n \simeq \Delta E_C$$

Thus we see that the _effective_ thermionic-field emission velocity decreases with increasing discontinuity in the conduction band. It is not immediately obvious why this velocity should depend on the spike height. The reason for the decrease in v_{tf} has, in fact, nothing to do with a decrease in the velocity of electrons; it has to do with the fact that as ΔE_C increases, the number of electrons at the top of the spike decreases, thus decreasing the electron flux across the interface, and

it is this decrease in flux that appears as an apparent decrease in the velocity in the final expression. The velocity of electrons remains v_x.

From the point of view of transistor performance, a large ΔE_C would make v_{tf} small, hence D large and thus J_{NC} small. Thus, the current gain will reduce with a larger ΔE_C. It is for this reason that it has been argued that HBTs be fabricated with a graded-gap emitter so as to reduce ΔE_C and thus attain large current gains. But there is an argument against grading, that if ΔE_C is large then the electrons entering the smaller band-gap base from a high emitter notch will be hot and therefore travel much faster across the base, thus decreasing the base transport factor and yielding a higher β [Enq87]. We shall look at this in more detail in Sec. 4.6.

Going back to the paper, the authors write eqn. (4.3-1) as

$$J_C = A_{21} \ [\exp(qV_{BE}/kT)-1] + A_{22} \ [\exp(qV_{BC}/kT)-1]$$

and similarly the emitter current as

$$J_E = A_{11} \ [\exp(qV_{BE}/kT)-1] + A_{12} \ [\exp(qV_{BC}/kT)-1] - J_{RG}^E$$

where J_{RG}^E is the recombination/generation current density in the base-emitter space-charge region. The above represent the Ebers-Moll-like equations for the HBT. The β of the transistor is derived as

$$\beta = \left[\ \cosh \ (W_B/L_1) \ - \ 1 + \frac{R_n J_{pE}}{R_p J_{nE}} \ \cosh(W_E/L_2) + \frac{J_{RG}^E \ R_n}{J_{nE} \ \exp(qV_{BE}/kT)} \ \right]^{-1}$$

where

$$R_n = 1 + \eta_n \ \cosh(W_B/L_1) \ \exp(-\Delta E_n/kT)$$

and

$$J_{nE} = \frac{q \ D_{n1} \ n_1}{L_1 \ \sinh(W_B/L_1)}$$

The effect of the wide gap emitter and the conduction band notch lies in the term $(R_n J_{pE}/R_p J_{nE})$ which is proportional to $\exp(\Delta E_G/kT)$, and also involves the ratio (v_d/v_{tf}).

The paper also has a section on extending the model to graded gap emitter HBTs wherein the grading is assumed to be such that it merely reduces the conduction band spike (thus increasing β) and does not affect anything else. The authors also report the derivation of the tunneling coefficient γ_n for abrupt as well as graded gap junctions.

Although some experimental data are provided, they are minimal and not sufficient to validate the model. Nonetheless, this model is more comprehensive than the one we discussed in the previous section.

4.4 An Ebers-Moll-Like Model

The fundamental difference between modeling an HBT and a homojunction bipolar transistor is that, in the former, the presence of the heterojunction must somehow be taken into account in deriving the I-V characteristics. Mathematically, this requires that the law of the junction, which for a homojunction is the well-known relation

$$pn = n_1^2 \ \exp(qV_{BE}/kT)$$

for the base-emitter junction (under low-injection forward bias), must be appropriately modified to take care of all the different current flow mechanisms that occur in a heterojunction. (For a reader interested in studying the boundary condition in a homojunction, the following are some recent papers on the subject: [Hau71, Hea79, Marsh80, For86]).

In the model that we now discuss [Lun86], Lundstrom introduces the concept of a generalized interface carrier velocity (S). The velocity assumes different expressions for different mechanisms of carrier transport across the heterojunction. We reproduce Lundstrom's derivation here, as we will be utilizing his formulation in chapters 5 and 6.

To begin with, the carrier flow across the heterojunction interface is assumed to be by thermionic emission. Thus the electron flux, F_{EN} (subscript E for emitter, N for electrons), across the emitter-base interface (at x=0 in Fig. 4.1) is written as the difference of the thermionic electron fluxes in the positive (F_1) and negative (F_2) directions:

$$F_N = F_1 - F_2$$
$$= v_x\, n(x=0^-) - v_x\, n(x=0^+)\, \exp(-\Delta E_C/kT) \qquad (4.4\text{-}1)$$

where v_x is the component of the electron thermal velocity in the x-direction. It is assumed that the electron quasi-Fermi level is nearly constant across the space charge layer except for a discontinuity at the interface. Lundstrom, in his paper, does not state this assumption, but it is implicit in his derivation. We shall discuss this assumption in chapter 5 (Sec. 5.2). This assumption allows one to write

$$F_{EN} = v_x\, n(x_{NE})\, \exp(-qV_{JN}/kT) - v_x\, n(x_{pE})\, \exp[(qV_{JP}-\Delta E_C)/kT] \quad (4.4\text{-}2)$$

In equilibrium, $F_{EN} = 0$, so that

$$n_0(X_{NE}) = n_0(x_{pE})\, \exp[(qV_{J0}-\Delta E_C)/kT] \qquad (4.4\text{-}3)$$

where the subscript 0 indicates equilibrium; thus V_{J0} is the equilibrium value of V_J, where $V_J = V_{JN} + V_{JP}$ (Fig. 4.1) is the electrostatic

potential difference across the base-emitter junction. Substituting Eqn. (4.4-3) into (4.4-2) and rearranging gives

$$F_{EN} = v_x \exp[(qV_{JP}-\Delta E_C)/kT] \ \{n_0(X_{pE})\exp[q(V_{J0}-V_J)/kT] - n(X_{pE})\}$$

In deriving the last expression, low injection is assumed in the emitter, so that $n(X_{pE}) \simeq n_0(X_{pE})$. Now defining a velocity, S_{EN},

$$S_{EN} \overset{\Delta}{=} v_x \exp \ [(qV_{JP}-\Delta E_C)/kT] \qquad (4.4-4)$$

and noting that $V_{J0}-V_J = V_{BE}$ if resistive drops are neglected, we get

$$F_{EN} = S_{EN} \ [\ n_0(X_{pE}) \ \{ \ \exp(qV_{BE}/kT) - 1 \ ] - \Delta n(X_{pE}) \ \} \quad (4.4-5)$$

Lundstrom calls S_{EN} the junction velocity. Note that the exponential factor in S_{EN} is due to the electron concentration decreasing from $x=X_{pE}$ to $x=0$ and has nothing to do with the magnitude of the electron velocity (as we discussed in the previous section). Lundstrom states that although carrier transport by thermionic emission was assumed in the above derivation, the boundary condition derived will be valid in general, for carrier transport by thermionic emission-tunneling, just thermionic emission, or drift-diffusion. Although Lundstrom does not provide a proof for his assertion, we can see from our discussion of Sec. 3.4, specifically Eqns. (3.4-1) through (3.4-4), that his assertion is, indeed, valid; because as we saw in that section, carrier transport across an interface can be written in the form of Eqn. (4.4-1) with the velocity given by Eqn. (3.4-4), which is a general expression for the velocity encompassing all possible transport mechanisms. Eqn. (4.4-5) thus is a generalized boundary condition at the heterojunction. Note that if S_{EN} is large, then $F_{EN}/S_{EN} \simeq 0$ and the equation reduces to the boundary condition for a homojunction.

Continuing with the derivation of the HBT model now, Lundstrom solves for carrier transport in the base by the Ebers-Moll procedure, to yield the electron flux through the base, F_B, neglecting recombination, as

$$F_B = D_n/W_B \ [\ \Delta n(X_{pE}) - \Delta n(X_{pC}) \ ] \qquad (4.4-6)$$

Noting that $F_B = F_{EN} = F_{CN}$, where F_{CN} is the electron flux across the base-collector heterojunction, one gets,

$$F_B = F_{EN} = F_{CN} = \frac{D_n n_{0B}/W_B}{1 + \dfrac{D_n/W_B}{S_{EN}} + \dfrac{D_n/W_B}{S_{CN}}} \ [\ \exp(qV_{BE}/kT) - \exp(qV_{BC}/kT) \ ]$$

$$(4.4-7)$$

Eqn. (4.4-7) is the major result of Lundstrom's paper. Note that it differs from a similar expression for a homojunction transistor in that it has two additional terms in the denominator of the pre-exponential factor, contributed from carrier transport across the two heterojunction interfaces. If both of these were homojunctions, then S_{EN}, $S_{CN} \gg$ (D_n/W_B), and the equation will reduce to its homojunction form.

Lundstrom then derives expressions for the various hole currents, and also extends the formulation to include recombination and energy gap grading in the quasi-neutral base.

The most significant aspect of this paper is the utilization of the concept of the junction velocity. As we discussed above, this concept makes this model comprehensive from the point of view of accounting for carrier transport across a heterojunction interface. This model, therefore, is a generalization of the model we looked at in the

previous section. Both of the models still do have the same limitations that the Ebers-Moll model has, which we discuss below.

4.5 A Gummel-Poon-Like Model

In deriving the basic bipolar transistor model, there are two main steps involved:

(1) Determining either the excess minority carrier density, or the pn product, on the base side of the space charge layer edge, as a function of the applied voltage. In homojunction transistors, the law of the junction determines these quantities, while in a heterojunction, as we saw in the previous two sections, the determination of this boundary condition is a more involved task.

(2) Solving for minority carrier transport in the base to obtain an expression for the current as a function of the carrier densities. The substitution of the result obtained in (1) then yields the I-V characteristics. The methodologies of Ebers-Moll [Ebe54] and Gummel-Poon [Gum70b] are the two major approaches used to solve for base transport, both for homojunction and heterojunction bipolar transistors. Of the two, the Ebers-Moll method yields a model that can only describe a transistor with a constant doping density and low injection in the base; and it also ignores the Early effect. The Gummel-Poon model, on the other hand, is applicable for any arbitrary doping profile and any level of injection in the base, and it also accounts for the Early effect.

The models we discussed in Secs. 4.3 and 4.4 used the Ebers-Moll formulation for step (2), as we saw. In this section, we discuss a model

[Ryu90] that takes a somewhat peculiar approach. It divides step (2) into two separate steps. Thus the approach is as follows:

(1) The excess electron density is determined the same way as Lundstrom did. Thus Eqn. (4.4-5) is used to determine Δn:

$$F_{EN} = S_{EN} \, [\, n_0(X_{pE}) \, \{ \, \exp(qV_{BE}/kT) - 1 \,] - \Delta n(X_{pE}) \, \} \quad (4.4\text{-}5)$$

(2a) Since Δn is still a function of the electron current (through F_{EN}: $F_{EN} = - J_N/q$), it is determined explicitly by solving the continuity equation in the base and equating F_{EN} with the current density obtained from that solution. The continuity equation is solved assuming low injection and zero electric field (hence no doping density gradient) in the base. Note that this is exactly what Lundstrom did (Eqn. (4.4-6) is the solution of the continuity equation in the base). Thus the method in this step of the model is Ebers-Moll-like, and thus suffers from the limitations of that model. Δn is thus found as

$$\Delta n(X_{pE}) = n_0(X_{pE})/\theta_n \, \{ \, [\exp(qV_{BE}/kT)-1] + \eta_n \, [\exp(qV_{BC}/kT)-1] \, \} .. (4.5\text{-}1)$$

where the subscript 0 refers to equilibrium,

$$\eta_n = \frac{(D_n/L_n)}{S_{EN} \, \sinh(W_B/L_n)}$$

and

$$\theta_n = 1 + \frac{(D_n/L_n)}{S_{EN}} \, \coth(W_B/L_n)$$

(2b) Now the carrier transport in the base is solved for again, this time using the Gummel-Poon methodology. The expression for the current density so obtained is a function of the pn product at the space charge layer edges. For n, Eqn. (4.5-1) is used. For p, it is assumed that $p \simeq$

p_0. Thus once again, low injection is invoked. The final expression for the current is obtained as

$$I_{CC} = -\frac{q^2 D_n n_{iB}^2 A^2 \{(1/\theta_n)[\exp(qV_{BE}/kT)-1] - (1-\eta_n/\theta_n)[\exp(qV_{BC}/kT)-1]\}}{qA \int_{X_{pE}}^{X_{pC}} p \; dx}$$

where A is the area of the device and I_{CC} is the current flowing through the base neglecting recombination. The expression has the same form as the Gummel-Poon expression, but since Δn is determined as described in step (2a) above, the model is applicable only under the conditions of the applicability of the Ebers-Moll model. Thus although ostensibly a Gummel-Poon-like model, it is not an improvement over Lundstrom's Ebers-Moll-like model.

Having derived an expression for I_{CC}, the authors derive expressions for the hole currents in the collector and emitter, and for recombination currents in the bulk of the emitter-base space charge region and at the heterojunction interface. They also have a section on base pushout and its effects on the various transit times. But it is not clear how these are incorporated into the model for J_C, since the transit times are functions of J_C themselves, and trying to include them in the model for J_C will give unwieldly transcendental equations. Also, the peculiar nature of the velocity vs. electric field characteristics in GaAs is not taken into account in modeling base pushout (but rather the model used for silicon is borrowed from [Bow73]). This, as we shall see in chapter 6, is erroneous.

4.6 Hot Electron Transport in the Base of HBT's

In Sec. 4.3 we mentioned in passing that in an (N/p/n) abrupt junction HBT, electrons injected into the base may be hot due to the conduction band spike, and this may cause the β of an abrupt junction HBT to be larger than that of a graded junction HBT despite the fact that the injection efficiency of the former is smaller. The issue of hot electron transport in the base of AlGaAs/GaAs HBT's has been examined, both numerically (using Monte-Carlo simulation) and experimentally. In this section, we review the papers written on this subject.

In 1982, Ankri and Eastman [Ank82a] first proposed the ballistic HBT (BHBT) for improved speed and high frequency operation. The idea was simple: if the number of collisions in the base of the transistor is minimized, electrons will travel ballistically, thus reducing the base transit time and improving speed. Collisions could be reduced by (i) reducing the base thickness and (ii) by providing the electrons a high kinetic energy, because at high energies, the scattering rate in GaAs is small. (But for kinetic energies above about 0.3 eV, the scattering rate again becomes large because electrons are scattered from the Γ valley to the L valley). Thus by having an abrupt heterojunction with a discontinuity less than 0.3 eV, the electrons can be launched into the base with a high kinetic energy.

But, concomitantly, an abrupt junction will reduce the injection efficiency of the junction, as we have discussed in previous sections. Thus there will be a trade-off between speed and gain. Ankri and Eastman discussed the above two points in a little detail; but without any

experimental verification of the idea. Thus, at that point in time, it remained merely another idea for improving speed. Later, they reported the fabrication of an AlGaAs/GaAs HBT with a cut-off frequency of 15 GHz, in which electrons traveled near-ballistically through the base [Ank83a, Ank83b].

Then, proceeding chronologically, there were two papers which reported Monte Carlo simulations of electron transport in the base of HBTs [Tom84, Maz86a].

The first of these, by Tomizawa et al [Tom84], simulated HBTs with a compositionally graded base, which created a quasi-field of about 20 kV/cm. The base was 1000°A thick and was doped to 2 x 10^{18} /cm^3. Their Monte Carlo simulations showed that the average electron energy in the base increased from kT (0.026 eV) at the emitter end to greater than 0.3 eV at the collector end, due to the quasi-electric field; or in other words, electrons moved horizontally in the energy band diagram in the base. The authors concluded from this that the transport must be ballistic. We think that there was not sufficient information to arrive at this conclusion. The only conclusion that could be drawn was that the average electron energy was higher than the thermal energy. The electrons could just have been travelling hot, and not necessarily ballistically.

There were two other points in the paper that are worth mentioning: First, the authors also simulated an uniform base HBT, in which the transport would be by diffusion alone, and found that the average electron velocity in the base _was_ an order lower than in the

graded base case, as expected. Second, that in the graded base HBT, the low doped (10^{16} /cm^3) collector went into high injection for a current density of about 1 mA/μm^2 (=10^5 A/cm^2) thus indicating the beginnings of base pushout.

The second paper reporting Monte Carlo simulations was by Maziar et al [Maz86a]. They simulated two different GaAs transistor structures proposed previously in the literature, to investigate their efficacy in reducing the base transit time.

The first was what they called the ballistic base HBT (BBHBT). (This was the structure proposed by Ankri and Eastman [Ank82a]). It consisted of a 9kT (0.23 eV at room temperature) or a 6kT (0.156 eV) discontinuity in the conduction band (or a "ballistic ramp", as they called it) and a 1200°A thick base with N_A = 5 x 10^{18} /cm^3. The results of simulation showed that the average velocity of electrons in the base was larger for the 6kT ramp than for the 9kT one; the explanation suggested for this was as follows: In the 9kT device, more electrons started out in the L valley, thus there was more (intervalley) scattering in the base and hence a smaller average velocity. It was not clear from their results whether transport was ballistic. The results they reported showed electron velocity distributions as functions of velocity, for the 9kT and 6kT ramps, 20 A° from the emitter. The paper did not have any results showing the distribution at the collector-base interface, which would have indicated whether the transport was ballistic or not. But even at 20 A°, substantial scattering was seen to have occurred. At least from these results one would conclude that in

both 9kT and 6kT structures, substantial scattering did occur over the 1200 A° of base thickness. It is difficult to conclude though, whether the electrons got hot or not: they at least had a higher average velocity than for a structure without a ramp. The conclusion of the authors was that BBHBTs are effective in reducing the base transit time for narrow bases; and they gave optimization curves of base thickness versus ramp height for a given transit time, and transit time versus base thickness for a given ramp height.

The second transistor structure the authors examined was a graded base HBT (GBHBT) in which the base was compositionally graded to create a quasi-electric field that would accelerate the electrons in the base. Compositional grading requires increasing the Al fraction from zero at the base-collector junction to a certain value at the emitter-base junction to achieve a desired quasi-field. Then for maintaining a high injection efficiency, it becomes necessary to have the Al fraction in the emitter higher than its highest value in the base. This has the consequence that a large fraction of electrons injected from the emitter reside in the upper valleys, thus increasing intervalley scattering. Hence trying to increase the quasi-field beyond a certain limit only degrades the base transit time. The authors therefore suggested that a dopant density gradient be used instead, and this did yield a superior performance. Thus, the concept of a graded base HBT did not seem a particularly attractive one.

There was one more Monte Carlo study of transport in graded heterostructures (Al-Omar and Krusius [Oma87]), but it was done on

$N^+/N/N^+$ devices and had no information relevant to the HBT. We shall not discuss it here.

Then there were three papers reporting experimental studies of electron transport in the base of AlGaAs/GaAs HBTs [Hayes86, Enq87, Ram88].

The first of these [Hayes86] (by Hayes, Levi, Gossard and English) used a technique called hot electron spectroscopy, which the authors themselves had proposed earlier, to investigate the physics of hot electron transport in the base of an HBT. The technique was briefly as follows:

An abrupt heterojunction transistor is fabricated with about the same Al concentration in the collector and emitter. Forward biasing the emitter-base junction causes electrons to be injected into the base with a high initial kinetic energy (ϕ_{be}). On reaching the collector end, only those electrons which have sufficient energy to surmount the base collector barrier (ϕ_{bc}) contribute to the collector current. Now ϕ_{bc} can be varied by varying V_{BC}. Thus by analyzing the variation of the collector current with V_{BC}, the energy distribution of the electrons reaching the collector can be determined.

The device structures studied by the authors had base thicknesses of either 900 or 450°A, base doping of $N_A=2\times10^{18}$ /cm^3, and emitter and collector Al fractions of 0.31 and 0.35 respectively, both doped to 2×10^{17} /cm^3. From the experimental results, they concluded that electrons reaching the collector had less energy than the base-emitter barrier, indicating that the transport was not ballistic; but the

electrons had a much higher average kinetic energy than kT, indicating that they were hot. At a lattice temperature of 4.2 K, the authors found the effective electron temperatures at the collector to be 150 K and 500 K for the 900°A and 450°A base devices, respectively.

Two comments are in order. First, the method used to evaluate the energy spectrum of the electrons reaching the collector assumed that there was no electron injection from the collector into the base. This would not be valid when the base-collector junction was forward biased, as was required in the experiment. Thus, we have some reservations about the accuracy of the conclusions drawn. Second, it is likely that there was substantial intervalley scattering in the base which checked ballistic transport. According to at least one group [Oma87], an Al concentration of more than 23% would cause intervalley scattering to be significant. Thus the categorical statement by the authors that ballistic transport does not occur in any AlGaAs/GaAs HBTs is not correct. One must perform experiments on devices with lower Al concentrations before making such a statement.

The second experimental study of electron transport in HBTs was by Enquist, Ramberg and Eastman [Enq87]. They compared the characteristics of graded and abrupt junction HBTs and used electroluminescence measurements as an added means of comparison.

All their devices had an AlGaAs emitter with an Al fraction of 0.22, doped to 2×10^{17} /cm^3, a base 1000°A thick with various doping concentrations, and a collector doping concentration of 5×10^{16} /cm^3. The important results and conclusions were as follows: For all the devices

studied, the abrupt junction transistors had higher β's than their graded junction counterparts. The electroluminescence measurements showed that the radiation emitted was from the GaAs base, and the integrated intensity was higher in graded junction devices (indicating more recombination). Measurements of β as a function of temperature showed that in graded junction transistors, the β's tended to zero as T→0 K while in abrupt junction transistors, they approached a non-zero value. From the above three results, and other supporting measurements, the authors concluded the following: β in the HBTs studied was governed by the base transport factor, and not by injection efficiency. In abrupt junction HBTs, electrons injected into the base were hot, hence they traveled at high velocities, spent a smaller time in the base, and therefore experienced less recombination than in their graded junction counterparts.

The third paper on this subject was by Ramberg and Ishibashi [Ram88], which also used electroluminescence measurements to characterize electron transport in HBTs. It was a detailed paper reporting measurements on many different HBT structures, but the results on electron transport in the base of the transistor were few and not conclusive. Hence we shall not discuss the paper here any further.

The conclusions for this section are as follows: Both Monte Carlo simulations and experimental results indicate that electrons are hot in the base of an HBT if the emitter-base heterojunction is abrupt and if the base thickness is small (< 2000°A). This results in smaller base transit times, but may result in smaller injection efficiencies due to

the large barrier that the electrons must surmount at the interface. Thus there is a trade-off between gain and speed (which has not been modeled quantitatively yet).

4.7 Velocity Overshoot in GaAs HBT's

Velocity overshoot is a phenomenon where the carrier velocity is found to reach values many times larger than the maximum value it would attain in steady-state. Overshoot occurs in the presence of sharp spatial or temporal gradients in the electric field. A sharp spatial gradient in the electric field may be present at the base edge of the base-collector space charge region (SCR) in a bipolar transistor. This gradient can cause the velocity of the carriers entering the SCR to overshoot over a few mean free paths. Such overshoot will reduce the transit time of carriers across the SCR, and thus will affect the cut-off frequency (f_T) of the transistor (since the transit time across this SCR contributes a significant fraction of the total transit time across the transistor).

In this section, we will review the papers that have addressed the issue of velocity overshoot in GaAs HBT's through Monte Carlo simulations. We will discuss the physical reasons for overshoot and an attempt at analytical modeling in Sec. 10.3.

Velocity overshoot was first reported in the literature by Ruch [Ruc72], who did numerical simulation of short-channel field effect transistors. Velocity overshoot in n-GaAs has since been studied in detail [Malo77, Kra78, Coo82, Tei83]. Velocity overshoot in the base-collector SCR of GaAs HBT's, was first reported by Asbeck et al

[Asbe82]. The model they used was simplistic; they found that including overshoot increased the f_T of the transistor under study from 55 GHz to 73 GHz.

Maziar et al [Maz86b] did Monte Carlo simulations on carrier transport in GaAs HBT's. They found that although velocity overshoot did occur in the base-collector SCR, its effect on the total transit time -- and hence on f_T -- was not significant. From their results they realized that velocity overshoot could be sustained over a longer distance if the gradient of the electric field was made smaller at the base edge of the SCR. This was because velocity overshoot subsided as electrons transferred to the higher L and X valleys; thus if this transfer was slowed down overshoot will persist longer. To effect this they proposed a new transistor structure: N-AlGaAs / p-GaAs (N_A=10^{19}/cm^{-3}, 750 A° thick) / p-GaAs (N_A = 3.5x10^{16}cm^{-3}, 3000 A° thick) / n-GaAs (N_D=10^{19}cm^{-3}). Such a structure is different from a conventional N/p/n HBT in that the lightly doped n-type collector is replaced with a lightly doped p-type region followed by a heavily doped n-type region. Such a structure caused the electric field to vary more gradually at the base edge of the SCR. Monte Carlo simulations showed that in such a structure, velocity overshoot persisted over a much longer length within the SCR, and an improvemnt of 25 percent in the transit time across the SCR was observed, as compared to the conventional structure.

Yamauchi and Ishibashi [Yam86a] measured the variation of f_T as a function of the base-collector bias (V_{BC}) experimentally, and found a significant dependence of f_T on V_{BC}. They attributed it to velocity

overshoot in the SCR. At least one author [Roc88] has claimed that the f_T variation was more likely to be due to a parasitic capacitance rather than velocity overshoot.

Rockett [Roc88] did a detailed Monte Carlo study of the base-collector SCR carrier transport. His conclusions were the same as those of Maziar et al [Maz86b]. That velocity overshoot had a negligible effect on the transit time across the SCR, and hence on f_T. His results indicated that beyond a few hundred angstroms into the SCR, a majority of electrons resided in the higher (L and X) valleys. He further pointed out that these Monte Carlo results did not agree with the results of Asbeck et al [Asbe82] whose model, Rockett said, was too simplistic.

Morizuka et al [Mor88] fabricated two transistor structures, one a conventional HBT, and the other proposed by Maziar et al [Maz86b] that we mentioned above. They found that as predicted by the Monte Carlo simulations, the f_T performance of the newly proposed structure was superior to the conventional HBT. They moreover found that even for the conventional HBT, the measured f_T at high current densities was larger than that predicted by simple calculations. They attributed this behavior to the accumulation of free carriers within the SCR which caused the electric field gradient to be less sharp.

Hu et al [Hu89] did a transient analysis on various HBT structures using Monte Carlo simulations. They also found that $N/p^+/p^-/n$ and $N/p^+/i/p^+/n$ structures had smaller turn-off times than the conventional $N/p^+/n$ structure.

Horio et al [Hor89] did numerical simulations on GaAs HBT's with a graded energy gap in the base. Their simulator accounted for hot electron transport as well as overshoot phenomena. Their results showed that in a graded-base HBT, electrons were hot in the base; and velocity overshoot occurred in the base-collector SCR. A comparison of the f_T values calculated with and without including the high energy effects showed that f_T's calculated with these effects included were substantially higher than those calculated without including these effects. The paper though did not separately consider the effects of hot electrons in the base and velocity overshoot in the SCR. It was therefore not clear how important overshoot was in determining f_T.

Katoh et al [Kat89b] used a self-consistent particle simulator to study GaAs HBT's and found, like the others, that a p-type collector reduced the transit time across the base-collector SCR due to overshoot effects.

In conclusion, we observe three points.

(i) The transit time across the base-collector SCR forms a significant fraction of the total emitter-to-collector transit time in GaAs HBT's.

(ii) In conventional N/p/n GaAs HBT's, velocity overshoot is observed in the base-collector SCR; but the effect of this overshoot on the transit time is negligible, since overshoot persists over a very short distance within the SCR.

(iii) By devising alternate structures that would make the gradient of the electric field less sharp, velocity overshoot can be made to persist

over a longer distance; in such structures, significant reductions in the transit time can be achieved.

4.8 GaAs HBT Technology

The following is a brief review of the developments in the fabrication technology of (primarily GaAs) HBTs.

The first HBT fabricated was perhaps in 1958 [Jen58]. Phosphorous was diffused into GaAs to from a GaAlP/GaAs HBT. The current gain (β) of the transistor was 0.9, and its low value was attributed to large contact resistances. Then with the advent of vapor phase epitaxy (VPE) [Rut60], more HBTs -- Ge/GaAs [Jad69], ZnSe/Ge [Hov69] -- were fabricated. β's upto 60 [Hov69] were measured on N-ZnSe/p-Ge/n-Ge HBTs; the major source of base current was found to be tunneling-interface recombination (see Fig. 3.1(c)). In the seventies and early eighties, liquid phase epitaxy (LPE) was utilized to yield better quality GaAs HBTs [Dum72, Alf73, Kon75, Ros77, Marty79, Bai80, Ank82b]. β's as high as 2000 and cut-off frequencies (f_T) as high as 5 GHz were achieved [Ank82b]. The transistor areas were large and were primarily responsible for limiting f_T. It is of interest to note here that there has been no remarkable increases in the current gain of HBTs in the last ten years. The efforts seem to have focused on improving the high frequency performance (with microwave applications in view). In fact, most of the high speed transistors reported in the literature have had low current gains (< 50). This is attributed to the small sizes of transistors required to attain such high speeds, which cause recombination at the surface [Henr78a, Henr78b, Hay87, Lyo89] and at the base contact

[Nak85a, LeeW89] to dominate and limit the current gain. There has though been some effort to remove these two recombination sources limiting β [Nak85b, Tiwa87, Nott88], and work continues to be done.

Most modern HBTs are fabricated using molecular beam epitaxy (MBE) or metal-organic chemical vapor deposition (MOCVD). There has been a steady improvement in the cut-off frequency. An f_T of 90 GHz was reported [Nag89] on an AlGaAs/GaAs HBT. InGaAs/InP HBTs utilizing hot electron transport in the base have yielded f_T's as high as 110 GHz [Not89]. Table 4.1 summarizes the history of the developments in GaAs HBTs.

Reference	Process	β	f_T(GHz)
Dum72	LPE	25	--
Kon75	LPE	350	--
Ros77	LPE	1800	0.4
Bai80	LPE	200	1
Ank82	LPE	2000	5
Asbe82	MBE	40	10
Mil83	MBE	--	16
Ito84	MBE	230	25
Rez86	MBE	3000	3
Fis86	MBE	--	35
Nak86	MBE	--	45
Cha87	MBE	28	55
San87	MOCVD	80	--
Nag89	MBE	25	80
Ma89	MBE	--	90

TABLE 4.1 History of Technological Developments of GaAs HBT's

CHAPTER 5
A NEW CHARGE-CONTROL MODEL FOR HBT'S

5.1 Introduction

In Secs. 4.2 through 4.5 in the last chapter, we discussed four major analytical models proposed for describing the dc characteristics of HBTs. Of these four models, two were the Ebers-Moll type, and two were the Gummel-Poon type. We also saw that of the two Gummel-Poon-like models, the first (described in Sec. 4.2) had little analytical merit, and the second (by Ryum and Abdel-Motaleb [Ryu90] described in Sec. 4.5) was strictly valid only under the conditions that the Ebers-Moll-like models were valid. Now one major reason why the Gummel-Poon model is so widely used in homojunction bipolar transistor circuit simulation is that it involves a <u>single</u> charge-control expression that accounts for arbitrary doping profiles and all levels of injection in the base. The model of Ryum and Abdel-Motaleb did not share this advantage. In this chapter, therefore, we derive a new charge-control model for the heterojunction bipolar transistor that overcomes the ambiguity of Ryum and Abdel-Motaleb's model that we looked at in Sec. 4.5.

During the course of deriving this model, we will discover a limitation of the Gummel-Poon model that seems to have gone unnoticed ever since the model was proposed: The Gummel-Poon model is found to be invalid if, when the transistor is operating in saturation (both junctions forward biased), one of the junction biases causes moderate or high injection in the base (see Sec. 5.3.4 for the relevant discussion).

This chapter is organized as follows. In Sec. 5.2 we list the assumptions made in deriving the model; in Sec. 5.3, we derive the new model for the collector current, and also discuss briefly the various components of the base current; in Sec. 5.4 we look at large and small signal equivalent circuits for the HBT; in Sec. 5.5 we present some numerical results based on this new model, and conclude the chapter in Sec. 5.6.

5.2 List of Assumptions

Throughout our treatment, we will assume an N-AlGaAs/p-GaAs/N-AlGaAs double, abrupt heterojunction transistor, although the methodology will be general enough to be applicable to any heterojunction system. The band diagram for such a transistor is illustrated in Fig. 5.1 (see Sec. 4.1 for a description of this figure). In the figure, we show a discontinuity in the electron quasi-Fermi level at the heterojunction interface; such a discontinuity is a consequence of carrier flow being by thermionic emission across this interface [Per64]. (This will become clearer when we derive the boundary conditions in the next section. Note that carrier transport by thermionic emission invalidates the conventional expression for the current density, $J_N = \mu_n\, n\, \nabla E_{FN}$). Although both junctions are shown abrupt in the figure, the model derived is applicable equally to both abrupt and graded junctions. (An "abrupt" heterojunction is one in which no energy gap grading is introduced intentionally. Abrupt junctions fabricated using MBE or MOCVD have interface thicknesses of 20-50 A° [Gar79a, Gar79b]).
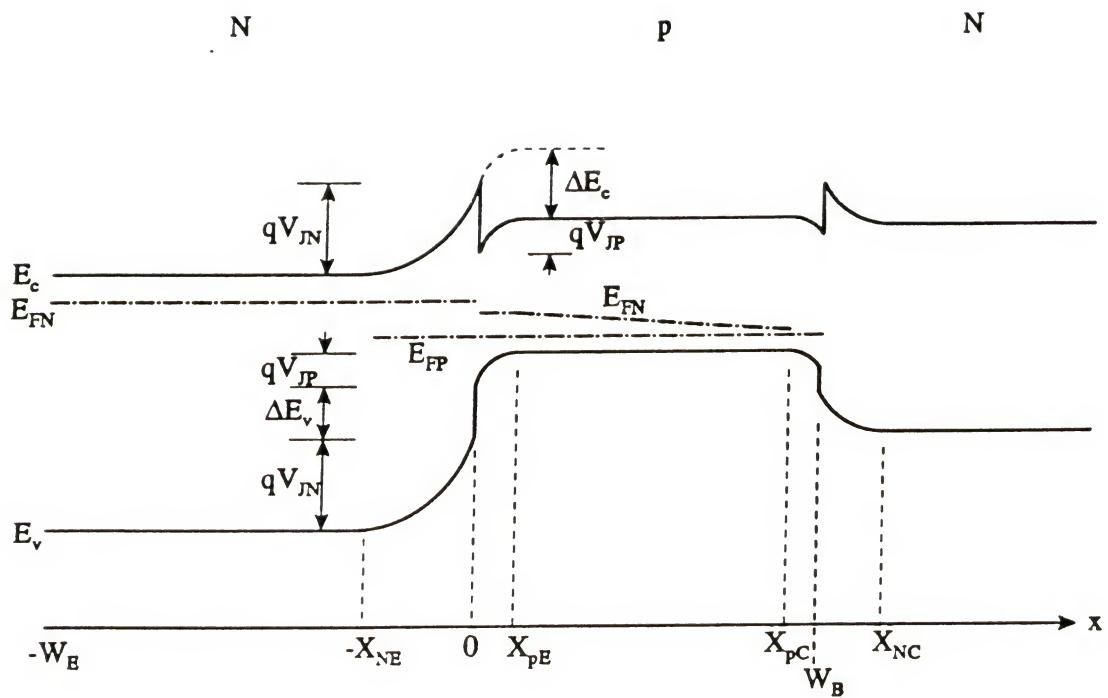
Fig. 5.1 Energy band diagram of an abrupt double heterojunction bipolar transistor.

Below we list the assumptions used in deriving the model.

(1) Hot electron transport in an abrupt junction transistor is negligible; such transport may be significant [Tom84, Maz86, Hayes86, Enq87, Ram88] depending on the height of the conduction band spike and the thickness of the base.

(2) In the base, the gradient of E_{FP} (majority hole quasi-Fermi level) is much smaller than that of E_{FN} (minority electron quasi-Fermi level); that is

$dE_{FP}/dx \ll dE_{FN}/dx$

The assumption is valid as long as $|J_P| \ll |J_N|$, since in the base of an N/p/N transistor, $p \geq n$ for any injection level. The assumption therefore will not be valid when β (current gain) < 1, that is, in deep saturation or at low current levels (when the space charge recombination current is comparable to the collector current).

(3) The intrinsic carrier density in the base (n_{iB}) is not a function of x; or in other words, the energy gap is constant in the base, and the device is isothermal.

(4) The quasi-Fermi levels (QFLs) are nearly flat across the space charge regions except for a possible discontinuity in the electron QFL at the heterojunction interface (such a discontinuity, as we mentioned earlier, will be present if carrier flow across the interface is by thermionic emission, as discussed in Ref. [Per64]). Although this assumption is not valid under reverse bias, since carrier densities in the space charge region are extremely small under reverse bias, the

position dependence of quasi-Fermi levels does not affect the model for the terminal characteristics.

(5) The recombination current in the quasineutral base can be neglected while deriving the charge-control relation for electrons. It is assumed that the recombination current can be algebraically added to this charge-control relation (derived in the next section) to obtain the terminal currents (see Gummel [Gum70a] for a detailed treatment of this issue).

The space charge region recombination current is also considered negligible. We will derive a model for that current and incorporate it into the charge-control model in the next chapter.

(6) There is no spike in the valence band at the heterojunction interface. This restricts the formulation presented below to heterojunctions that have only a conduction band spike; but the methodology proposed can still be used to include spikes in the valence band.

(7) The intrinsic carrier densities on the two sides of the junction (from the emitter to the base for example) can be related as

$$n_{iE}^2 \, \exp(\Delta E_G/kT) \simeq n_{iB}^2,$$

where ΔE_G is the difference in the energy gaps of the two sides. This neglects the changes in the density of states N_c and N_v, since

$$n_i^2 = N_C \, N_v \, \exp \, (-E_G/kT)$$

(8) Boltzmann statistics are valid.

(9) The treatment throughout is one dimensional. Thus all two or three dimensional effects are ignored.

(10) DC steady state is assumed in the analysis of Sec. 5.3.

5.3 A New Charge-Control Model

Our methodology will be to first derive Gummel's charge-control relation [Gum70a]; then determine the pn products at the heterojunction space charge layer boundaries in the base, and insert these into the charge-control relation to solve for the current density; and then evaluate the various terms occurring in the charge-control relation. We note here that the major difference in charge-control models for the homojunction and heterojunction transistors is in the determination of the pn product at the space charge layer boundary. This difference, in turn, is due to the additional current mechanisms (thermionic emission, tunneling etc.) occurring in an HBT at the heterojunction interface, as opposed to the simpler drift-diffusion flow in a homojunction transistor.

5.3.1 Gummel's Charge-Control Relation

To derive Gummel's charge-control relation [Gum70a] we begin with the current density equations in the base of the transistor:

$$J_N = \mu_n \ n \ dE_{FN}/dx$$

$$J_P = \mu_p \ p \ dE_{FP}/dx$$

Assumption (2) allows us to write

$$J_N \simeq \mu_n \ n \ d(E_{FN}-E_{FP})/dx \tag{5.3-1}$$

Now differentiating the relation

$$np = n_{iB}^2 \ \exp \ [(E_{FN}-E_{FP})/kT]$$

and substituting in (5.3-1) gives

$$J_N = \mu_n \ kT \ \frac{n_{iB}^2}{p} \ \frac{d}{dx} \left(\frac{pn}{n_{iB}^2} \right) \tag{5.3-2}$$

where n_{iB}^2 is the intrinsic carrier density in the base. Rearrangement, Einstein's relation, assumption (5), and integration across the quasineutral base region, from $x=X_{pE}$ to X_{pC}, yields [Kro85b],

$$J_N = q \frac{(pn/n_{iB}^2)_{x=X_{pC}} - (pn/n_{iB}^2)_{x=X_{pE}}}{\int_{X_{pE}}^{X_{pC}} (p/D_n n_{iB}^2) \, dx} \tag{5.3-3a}$$

where X_{pC} and X_{pE} are defined in Fig. 5.1. <u>The above equation is a general charge-control relation for any bipolar transistor</u>; it is valid even for transistors with a graded energy gap in the base, as was discussed by Kroemer [Kro85b]. To obtain a closed-form analytical model though, we assume that the energy gap is constant in the base, so that n_{iB} is not a function of x; and then use the mean value theorem of calculus to write the denominator as

$$\int_{X_{pE}}^{X_{pC}} (p/D_n n_{iB}^2) \, dx = \frac{1}{\overline{D_n} n_{iB}^2} \int_{X_{pE}}^{X_{pC}} p \, dx$$

where $\overline{D_n}$ is the average value of D_n over the quasineutral base. In the final charge-control expression, D_n gets absorbed into the pre-exponential saturation current I_S, which is determined by terminal measurements [Get78]. Since D_n is in general a function of the doping density, this formulation maintains the validity of the model even when the doping density in the base is not constant. For convenience, $\overline{D_n}$ is simply written as D_n henceforth. Eqn. (3a) then can be written as

$$J_N = q \, D_n \frac{(pn)_{X_{pC}} - (pn)_{X_{pE}}}{\int_{X_{pE}}^{X_{pC}} p \, dx} \tag{5.3-3b}$$

It is at this point that a model for a heterojunction transistor becomes different from that for a homojunction one. In the latter, the pn products at $x=X_{pC}$ and X_{pE} are evaluated by utilizing the law of the junction:

$$pn = n_i^2 \exp(qV_A/kT) \tag{5.3-4}$$

where V_A is the voltage applied across the junction. This relation assumes that the quasi-Fermi levels for electrons and holes remain nearly constant across the space charge layer of the junction. The near-flatness of the quasi-Fermi levels in turn is caused by the fact that in the space charge layer of a homojunction, the drift and diffusion components of carrier flow (which make up the two components of $J_N = \mu_n$ $n\, dE_{FN}/dx$, for electrons) are individually much larger in magnitude than the net current flowing through the layer; thus the net current flowing is caused by a minor perturbation of the opposing drift and diffusion tendencies. In a heterojunction, on the other hand, the presence of a spike in the energy band introduces the additional carrier flow mechanisms of thermionic emission and tunneling [Gri84], which now dominate the flow across the heterojunction interface; this in turn causes the quasi-Fermi level to have a discontinuity at the interface [Per64]. Hence it is no longer possible to determine the pn product from Eqn. (4). To determine these pn products, we must analyze the transport of carriers across the heterojunction interface. For this we utilize Lundstrom's concept of the junction velocity [Lun86] that we discussed in Sec. 4.4.

5.3.2 Boundary Conditions at a Heterojunction

The first three steps of the derivation below are identical to the ones we looked at in Sec. 4.4. They are repeated here for the sake of completeness. Since we already discussed the terminology and the concept of the junction velocity in Sec. 4.4, we will not repeat these here. The derivation below is different from the one in Sec. 4.4 (this latter was, of course, Lundstrom's derivation) in that it is not restricted to low injection levels in the emitter.

The electron flux (F_{EN}) across the emitter-base interface (at x=0 in Fig. 5.1) is written as the difference of the electron fluxes (F_1 and F_2) in the positive and negative x directions:

$$F_{EN} = F_1 - F_2$$
$$= v_x \, n(x=0^-) - v_x \, n(x=0^+) \, \exp(-\Delta E_c/kT) \qquad (5.3-5)$$

Assumption (4) then gives

$$F_{EN} = v_x \, n(X_{NE}) \, \exp(-qV_{JN}/kT) - v_x \, n(X_{pE}) \, \exp[(qV_{JP}-\Delta E_c)/kT]$$

where $V_J = V_{JN} + V_{JP}$ (see Fig. 5.1), is the electrostatic potential difference across the base-emitter space charge layer. Now since (see Eqn. 4.4-4))

$$S_{EN} \stackrel{\Delta}{=} v_x \, \exp\left[(qV_{JP}-\Delta E_c)/kT\right] \qquad (5.3-6)$$

we get

$$F_{EN} = S_{EN} \left\{ \, n(X_{NE}) \, \exp\left[(-qV_J+\Delta E_c)/kT\right] - n(X_{pE}) \, \right\} \qquad (5.3-7)$$

Rearranging gives

$$n(X_{pE}) + F_{EN} \, / \, S_{EN} = n(X_{NE}) \, \exp\left[(-qV_J+\Delta E_c)/kT)\right] \qquad (5.3-8)$$

Equation (5.3-8) is the boundary condition relating the electron densities on the two sides of a heterojunction. The second term on the

left hand side distinguishes this equation from that for a homojunction; this term represents the effect of the band spike, ΔE_C, on carrier densities. If the spike is small, S_{EN} is large, the (F_{EN}/S_{EN}) term is negligible compared to the other two terms, and thus the equation reduces to that for a homojunction (note also that if the spike is absent, $S_{EN} = v_x$).

A similar relation for holes can be obtained as

$$p(X_{pE}) + F_{EP} / S_{EP} = p(X_{NE}) \exp [(qV_J+\Delta E_v)/kT] \qquad (5.3\text{-}9)$$

Now there is no interface spike in the valence band of an N-AlGaAs/p-GaAs junction (see Fig. 5.1, and also assumption (6)). Thus hole flow across the junction is as in a homojunction, and we can neglect the (F_{EP}/S_{EP}) term (since S_{EP}, the effective interface hole velocity,, will be large compared to the hole diffusion velocity [Lun86]). Multiplying the last two equations then, and rearranging, yields

$$p(X_{pE})n(X_{pE}) = - (F_{EN}/S_{EN}) \; p(X_{pE}) + p(X_{NE})n(X_{NE}) \exp(\Delta E_G/kT) \qquad (5.3\text{-}10)$$

where $\Delta E_G = \Delta E_c + \Delta E_v$. Now since hole flow is as in a homojunction, we can assume that the hole quasi-Fermi level is constant across the space charge region. This allows us to write (see Fig. 5.1)

$$p(X_{NE}) \; n(X_{NE}) = n_{iE}^2 \exp \{ \; [\; E_{FN}(X_{NE}) - E_{FP}(X_{NE}) \;] \; / \; kT \; \}$$
$$= n_{iE}^2 \; \exp(qV_{BE}/kT)$$

Further, from assumption (7)

$$n_{iE}^2 \; \exp(\Delta E_G/kT) \simeq n_{iB}^2$$

Combining the last three equations, we obtain

$$p(X_{pE})n(X_{pE}) = - (F_{EN}/S_{EN}) \; p(X_{pE}) + n_{iB}^2 \; \exp(qV_{BE}/kT) \qquad (5.3\text{-}11)$$

Equation (5.3-11) defines the pn product at the space charge layer boundary of a heterojunction. The pn product at $x = X_{pC}$ is similarly

$$p(X_{pC})n(X_{pC}) = - (F_{CN}/S_{CN}) \; p(X_{pC}) + n_{iB}^2 \exp(qV_{BC}/kT) \qquad (5.3\text{-}12)$$

5.3.3 The New Charge-Control Relation for the HBT

Combining the last two equations with Eqn. (5.3-3b), noting that $J_N = - F_{EN}/q = F_{CN}/q$, and integrating over the cross-sectional area, A, of the transistor yields the final expression as

$$I_{CC} \overset{\Delta}{=} J_N \; A$$

$$= - q^2 \; D_n \; n_{iB}^2 \; A \; \frac{[\; \exp(qV_{BE}/kT) \; - \; \exp(qV_{BC}/kT) \;]}{q \displaystyle\int_{X_{pE}}^{X_{pC}} p \; dx + \frac{qD_n}{S_{EN}} p(X_{pE}) + \frac{qD_n}{S_{CN}} p(X_{pC})}$$

$$\overset{\Delta}{=} - I_S \; Q_{B0}/Q_B \; [\; \exp(qV_{BE}/kT) \; - \; \exp(qV_{BC}/kT) \;] \qquad (5.3\text{-}13)$$

$$\overset{\Delta}{=} I_F - I_R \qquad (5.3\text{-}13a)$$

where $I_S = (q^2 \; A^2 \; D_n \; n_{iB}^2 \; / \; Q_{B0})$, $Q_{B0} = Q_B(V_{BE}=V_{BC}=0)$, and Q_B is defined in Eqn. (5.3-14) below. Equation (5.3-13) is a major result. The expression is identical in form to Gummel's charge-control relation [Gum70a]. Only now Q_B does not have the physical significance of being the total hole charge in the base, since there are two additional terms contributing to it from the current flow across the interfaces:

$$Q_B \overset{\Delta}{=} q \; A \int_{X_{pE}}^{X_{pC}} p \; dx + \frac{qAD_n}{S_{EN}} p(X_{pE}) + \frac{qAD_n}{S_{CN}} p(X_{pC})$$

$$\overset{\Delta}{=} q \; A \int_{X_{pE}}^{X_{pC}} p \; dx + \quad Q_{SE} \quad + \quad Q_{SC} \qquad (5.3\text{-}14)$$

5.3.4 Evaluating Q_B

We split the integral of Eqn. (5.3-14) in the usual manner [Ryu90]:

$$qA \int_{X_{pE}}^{X_{pC}} p \; dx = qA \left[\int_{X_{pEO}}^{X_{pCO}} N_{AB} \; dx - \int_{X_{pEO}}^{X_{pE}} N_{AB} \; dx - \int_{X_{pC}}^{X_{pCO}} N_{AB} \; dx + \int_{X_{pE}}^{X_{pC}} (p-N_{AB}) \; dx \right]$$

$$\overset{\Delta}{=} \quad Q_{BO} \quad + \quad Q_E \quad + \quad Q_C \quad + \quad (Q_F + Q_R)$$

$$\dots (5.3\text{-}15)$$

where N_{AB} is the acceptor concentration in the base and the subscript 0 refers to equilibrium quantities.

We now consider how each of the charge components in Eqns. (5.3-14) and (5.3-15) can be determined.

Charges Q_F and Q_R. The last term on the right hand side of Eqn. (5.3-15), (Q_F+Q_R), is the excess hole charge in the quasineutral base. The charges Q_F and Q_R are the contributions to this charge due to V_{BE} and V_{BC}, respectively. As Gummel and Poon did implicitly in their analysis [Gum70b], we assume that these two contributions can be linearly superposed [Get78], so that we can write

$$qA \int_{X_{pE}}^{X_{pC}} (p-N_{AB}) \; dx = Q_F + Q_R \simeq qA \int_{X_{pE}}^{X_{pCO}} (p-N_{AB})_{V_{BC}=0} \; dx$$

$$+ \; qA \int_{X_{pEO}}^{X_{pC}} (p-N_{AB})_{V_{BE}=0} \; dx \quad (5.3\text{-}16)$$

Equation (5.3-16) assumes that the excess hole charge in the quasineutral base can be calculated by superposing the excess hole charges caused by the biases across each of the junctions, while the other junction is kept at zero bias. Such superposition is valid only if the continuity equation is linear under the operating conditions being considered. The continuity equation, in turn, is linear only if the entire base layer is either in low injection or high injection.

To relate Q_F and Q_R to measurable quantities, they are written as

$$Q_F = \tau_F I_F \qquad \text{and} \qquad Q_R = \tau_R I_R \qquad (5.3\text{-}17)$$

where τ_F and τ_R are, respectively, the forward and reverse transit times across the quasineutral base [Ryu90], and I_F and I_R are defined by Eqn. (5.3-13a). Now in Eqn. (5.3-17), Q_F is the excess hole charge in the base calculated with $V_{BC} = 0$. But I_F as defined by Eqn. (5.3-13a) is a function of Q_B, which is a function of V_{BC} through Q_C and Q_R. Hence once again the expression for Q_F (Q_R) is not accurate when Q_C (Q_E) or Q_R (Q_F) are large. The effect of Q_C on Q_B (and hence on Q_F) is secondary. On the other hand, Q_R can become a large quantity if the base-collector junction causes high injection in the base.

The combined effect of the assumptions necessary for the validity of Eqn.s (5.3-16) and (5.3-17) is to make the Gummel-Poon model (and hence the present model) invalid if the base goes into moderate or high injection while both the junctions are forward biased. The model remains valid though for all levels of injection if one of the junctions is reverse biased, because then, on the right hand side of Eqn. (5.3-16), the term corresponding to the reverse biased junction contributes a negligible charge, and thus the assumption of superposition is no longer necessary.

Charges Q_E and Q_C. By the definition of Q_E, as given by Eqn. (5.3-15),

$$Q_E = -qA \int_{X_{pE0}}^{X_{pE}} N_{AB} \, dx = qA \, N_{AB} \, (X_{pE0} - X_{pE}) \qquad (5.3\text{-}18)$$

If $N_{AB} \neq N_{AB}(x)$, then we evaluate this as

$$Q_E = qA \ N_{AB} \ (X_{pE0} - X_{pE})$$

By the depletion approximation, X_{pE} can be evaluated as

$$X_{pE} = \left[\frac{2\varepsilon_E \varepsilon_B N_{DE} \ (V_{biE} - V_{BE})}{qN_{AB} \ (\varepsilon_E N_{DE} + \varepsilon_B N_{AB})} \right]^{1/2} \qquad (5.3\text{-}19)$$

where ε_E and ε_B are the dielectric permittivities in the emitter and the base respectively, and N_{DE} is the doping density in the emitter. A more accurate expression can be found by integrating Poisson's equation once to find (dV/dx) and writing

$$X_{pE} = \int_{V_0}^{V_{JP}} dV \ / \ (dV/dx) \qquad (5.3\text{-}20)$$

where V is the electrostatic potential, and $V_0 = V(x=0)$. X_{pE} can be determined numerically from Eqn. (5.3-20). If N_{AB} is not independent of x, the evaluation of Q_E will be more involved.

Now the usual method for evaluating Q_E and Q_C in Eqn. (5.3-15) has been to write

$$Q_E = \int_0^{V_{BE}} C_{JE} \ dV \qquad (5.3\text{-}21)$$

where C_{JE} is the base-emitter junction capacitance. Then using the definition of C_{JE}

$$C_{JE} \overset{\Delta}{=} \frac{dQ_E}{dV_{BE}} = \frac{d}{dV_{BE}} \ [\ qA \int N_{AB} \ dx \ ] \qquad (5.3\text{-}22)$$

(assuming the depletion approximation to get the last equality). The above, it seems to this author, is a circular method; since first Q_E is evaluated, then it is differentiated to find C_{JE}, and then it is

integrated to give back Q_E! The differentiation-integration cycle is redundant.

   Charges Q_{SE} and Q_{SC}. In Eqn. (5.3-14), Q_{SE} and Q_{SC} are functions of $p(X_{pE})$ and $p(X_{pC})$ respectively. Now given the boundary condition of Eqn. (5.3-11) at a heterojunction, there is no straightforward way of evaluating the hole densities at the space charge region edges. We therefore make a simplifying assumption: For evaluating Q_{SE}, for example, we assume

$$p(X_{pE}) \simeq N_{AB}(X_{pE}) \qquad\qquad (5.3\text{-}23a)$$

The justification for the assumption is as follows. From Eqn. (5.3-14), Q_{SE} $(=qAD_np(X_{pE})/S_{EN})$ is significant compared to the integrated hole density only when S_{EN} is small, which happens when the conduction band spike (ΔE_c) is large compared to kT. Now a large conduction band spike impedes the injection of electrons into the base. Thus in a typical HBT, in which the base is more heavily doped than the emitter, the base will remain in low injection. Hence we can assume that the base remains in low injection when S_{EN} is small. If on the other hand, S_{EN} is large, Q_{SE} will be small compared to the integrated hole density, and then there will be no need to evaluate $p(X_{pE})$. Similarly, for evaluating Q_{SC} we assume

$$p(X_{pC}) \simeq N_{AB}(X_{pC}) \qquad\qquad (5.3\text{-}23b)$$

We realize that such an assumption does, theoretically, restrict the scope of this model, and it can no more claim to be as generally valid as the Gummel-Poon model for the homojunction transistor. From a practical viewpoint, however, the assumption is not restrictive for the

reason we stated above (namely, practical HBTs have the base more heavily doped than the emitter). Thus we finally write

$$Q_{SE} \simeq qAD_nN_{AB}(X_{pE}) \,/\, S_{EN} \qquad\qquad (5.3\text{-}24a)$$

and

$$Q_{CS} \simeq qAD_nN_{AB}(X_{pC}) \,/\, S_{CN} \qquad\qquad (5.3\text{-}24b)$$

Charge q_b: Now dividing Eqn. (5.3-14) by Q_{B0}, and substituting Eqn.s (5.3-15,17,24) into it, we get

$$q_b \overset{\Delta}{=} 1 + \frac{Q_E}{Q_{B0}} + \frac{Q_C}{Q_{B0}} + \frac{\tau_F I_F}{Q_{B0}} + \frac{\tau_R I_R}{Q_{B0}} + \frac{qAD_nN_{AB}(X_{pE})}{S_{EN}} + \frac{qD_nN_{AB}(X_{pC})}{S_{CN}} \qquad (5.3\text{-}25)$$

Then we write

$$q_b = q_1 + q_2 \,/\, q_b$$

where

$$q_1 = 1 + \frac{Q_E}{Q_{B0}} + \frac{Q_C}{Q_{B0}} + \frac{qD_nN_{AB}}{Q_{B0}} \left[\frac{1}{S_{EN}} + \frac{1}{S_{CN}} \right] \qquad (5.3\text{-}26)$$

and

$$q_2 = \frac{q^2 D_n n_{iB}^2}{Q_{b0}^2} \{ \tau_F [\exp(qV_{BE}/kT)-1] + \tau_R [\exp(qV_{BC}/kT)-1] \} \qquad (5.3\text{-}27)$$

where we have used Eqn. (5.3-13a) for I_F and I_R. Thus we get

$$q_b = (q_1/2) + [(q_1/2)^2 + q_2]^{1/2} \qquad\qquad (5.3\text{-}28)$$

Eqn.s (5.3-13) and (5.3-28) together give the final expression for I_{CC}. All of the parameters in this expression can be determined by measurements at the device terminals, except for S_{EN} and S_{CN}. These effective velocities can be determined from knowledge of the band structure, as expressed on Eqn. (5.3-6), neglecting the impact

generation rate. The components of the base current, I_B, are described below.

5.3.5 Base Currents in an HBT

We now briefly discuss the various components of the base current.

(i) Recombination current in the quasineutral base region (I_{RB}): This is written as

$$I_{RB} = qA \int_{X_{pE}}^{X_{pC}} (p-N_{AB}) \, dx \, / \, \tau_{rec} = (Q_F + Q_R) \, / \, \tau_{rec} \qquad (5.3\text{-}29)$$

$$\therefore I_{RB} = \tau_F/\tau_{rec} \, I_F + \tau_R/\tau_{rec} \, I_R \qquad (5.3\text{-}29a)$$

where τ_{rec} is the recombination lifetime in the base, τ_F and τ_R are the forward and reverse transit times across the base, and I_F and I_R are defined by Eqn. (5.3-13a). I_{RB} is the so-called ideal component of the base current. The nonideal component of base current arises from the recombination in the base-emitter space charge region, or at the extrinsic base interface.

(ii) Recombination in the space charge region (I_{SCR}): For a heterojunction, modeling of this current, and incorporating it into the model formulated above, is non-trivial. In fact, we discovered that this current affects the magnitude of the collector current also. We will derive a model for I_{SCR} and its effect on I_{CC} in the next chapter. We just quote the final results here. I_{SCR} is written in the form

$$I_{SCR} = I_{R1} \exp(qV_{BE}/2kT) + I_{R2} \exp[(E_{FN}(X_{pE})-E_{FP}(X_{pE}))/kT]$$

$$= I_{R1} \exp(qV_{BE}/2kT) + I_{R2} [\, p(X_{pE}) \, n(X_{pE}) \,]^{1/2} \, / \, n_{iB}$$

$$(5.3\text{-}30)$$

where the two terms respectively represent the recombination to the left and to the right of the conduction band spike in Fig. 5.1. The expressions for I_{R1} and I_{R2} are involved and can be found in chapter 6. The pn product in the second term is given as

$$[p(X_{pE}) n(X_{pE})]^{1/2} = n_{iB} \exp(qV_{BE}/2kT) \{ -RT \cdot r_s + [(RT \cdot r_s)^2 + 1]^{1/2}\}$$

$$(5.3-31)$$

where

$$RT = \frac{I_{R2} N_{AB}^2}{q S_{NE} n_{iB}^2 \exp(qV_{BE}/2kT)} \quad , \quad r_s = \frac{1 + D_n/(W_B' S_{NC})}{1 + D_n/W_B'(1/S_{NE}+1/S_{NC})} \quad (5.3-32)$$

where $W_B' = (X_{pC}-X_{pE})$. I_{CC} given by Eqn. (13) gets modified in the presence of I_{SCR} as follows:

$$I_{CC} = - \frac{I_S Q_{B0}}{Q_B} [\exp(qV_{BE}/kT) - \exp(qV_{BC}/kT] \cdot$$

$$\left[1 + \frac{RT^2 r_s}{2r_n} - \sqrt{\frac{RT^2 r_s}{r_n^2} \left(1 + \frac{RT^2 r_s}{4} \right)} \right] \quad (5.3-33)$$

where

$$r_n = 1 - \frac{\exp(qV_{BC}/kT)}{\exp(qV_{BE}/kT)} \quad (5.3-34)$$

(iii) <u>Surface recombination</u>: This is a strong function of the fabrication process used; we write this current simply as

$$I_{RS} = I_{RS0} \exp(qV_{BE}/n_{rs}kT)$$

where I_{RS0} and n_{rs} are parameters that characterize the current.

(iv) <u>Hole current in the collector</u>: This is obtained by standard diode theory as

$$I_{PC} = qAD_p n_{iC}^2/(N_{DC}W_C') \exp(qV_{BC}/kT) \quad (5.3-35)$$

in low injection; W_C' here will either be the collector layer thickness (W_C), or the hole diffusion length (L_p) or a function of them both, depending on the magnitudes of each. High injection in the collector will be accompanied by quasisaturation, which we have ignored here. Finally, from assumption (5), the collector current is written as

$$I_C = I_{CC} - I_{RB} \qquad\qquad (5.3\text{-}36)$$

The base current is the sum of all the recombination currents:

$$I_B = I_{RB} + I_{SCR} + I_{RS} + I_{PC} \qquad\qquad (5.3\text{-}37)$$

5.4 Circuit Models for the HBT

Fig. 5.2(a) shows the large signal circuit model for the HBT, that has the same topology as that implemented in SPICE for a homojunction bipolar transistor [Ant88]. The model proposed in Sec. 5.3 above provides us with a more accurate functional dependence of the circuit elements shown in the figure. The current sources I_C and I_B are defined by Eqn.s (5.3-36) and (5.3-37) respectively. R_C, R_B and R_E are extrinsic resistances associated with each of the three terminals. C_{BE} and C_{BC} are the capacitances associated with the base-emitter and the base-collector junctions respectively, and comprise the sum of the junction capacitance and the diffusion capacitance [Ant88].

Fig. 5.2(b) shows a small signal model for the HBT. The various elements have their usual meanings:

$$g_{mF} = dI_C/dV_{BE} \qquad\qquad\qquad g_\pi = dI_B/dV_{BE}$$

$$g_o = dI_C/dV_{BC} \qquad\qquad\qquad g_\mu = dI_B/dV_{BC}$$

$$g_{mR} = dI_E/dV_{BC}$$

$$C_\pi = g_{mF}\tau_F + C_{JE} \qquad\qquad\qquad C_\mu = g_{mR}\tau_R + C_{JC}$$
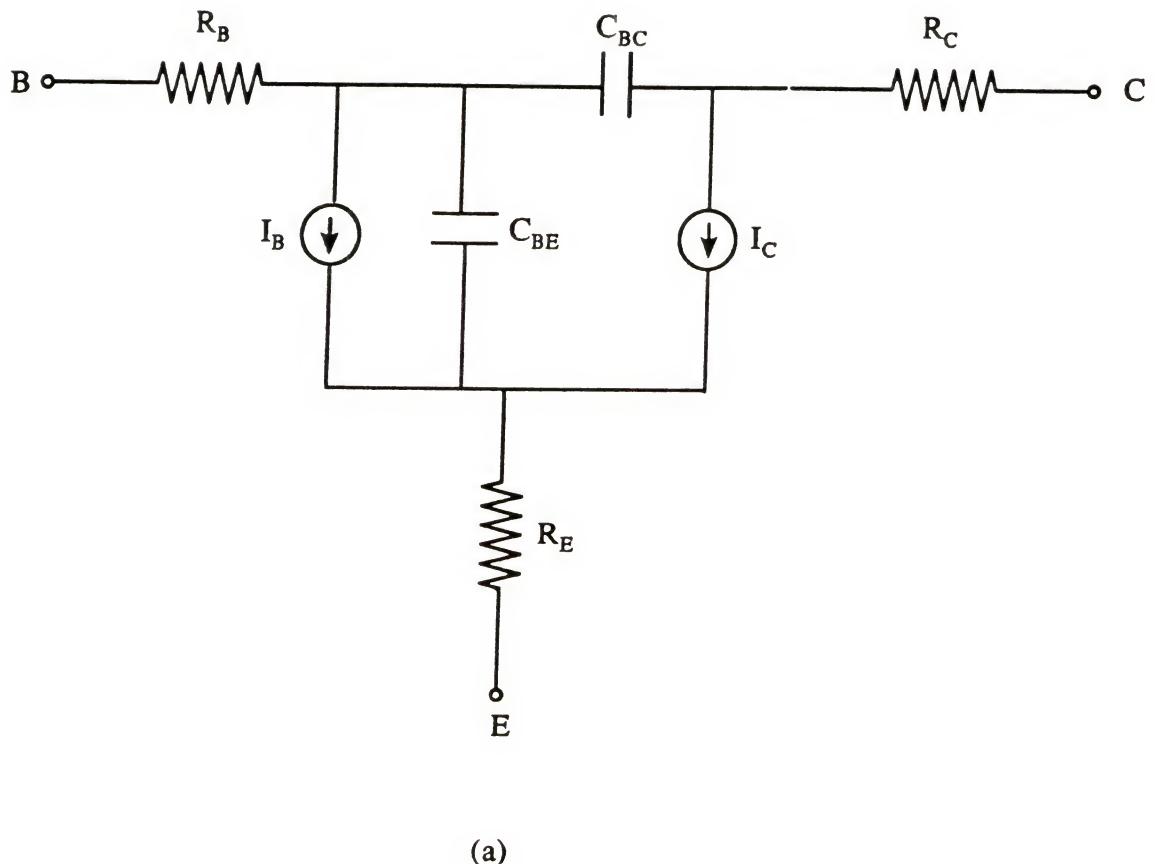
(a)

Fig. 5.2(a) A large-signal equivalent circuit model for the HBT.
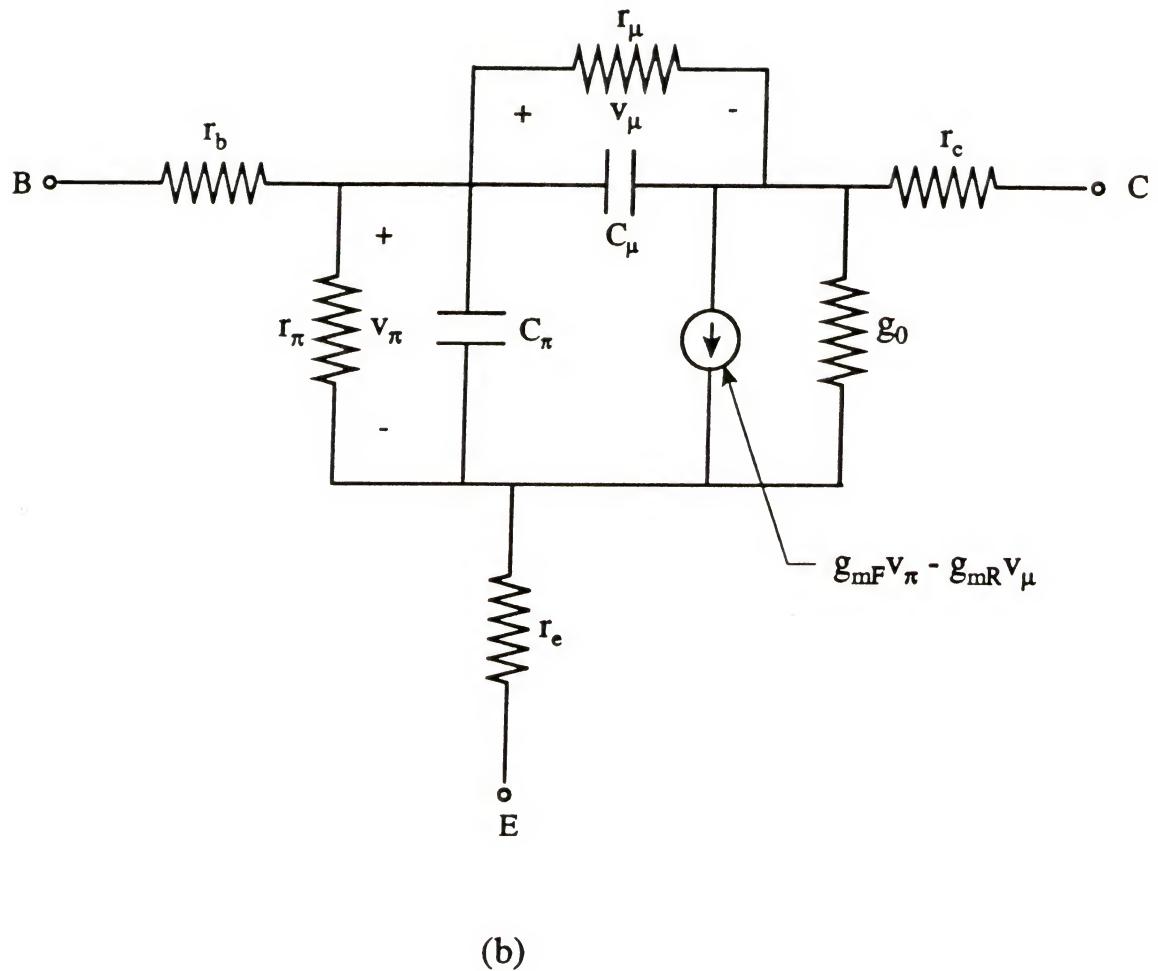
(b)

Fig. 5.2(b) A small-signal equivalent circuit model for the HBT.

where C_{JE} and C_{JC} are junction capacitances. The procedure for determining these parameters as functions of the terminal voltages and currents is the same as that used for homojunction transistors [Ant88]. The only difference is that for I_C and I_B, Eqn.s (5.3-36) and (5.3-37) must now be used.

5.5 Numerical Results

In Fig. 5.3 we compare the I_C-V_{BE} characteristics calculated from the present model with the numerical simulation done by Yokoyama et al [Yok84], for an abrupt N-$Al_xGa_{1-x}As$/p-GaAs/n-GaAs HBT. The various parameters used are: $\Delta E_G = 1.25x$, $\varepsilon/\varepsilon_0 = (13.1 - 3x)$, $\Delta E_C = 0.85\Delta E_G$, $\tau_p = \tau_n = 1$ ns (recombination lifetimes in the base), $\mu_n = 5000$ cm^2/V·s, $\mu_p = 100$ cm^2/V·s, $W_E = 0.25$ μm, $W_B = 0.05$ μm, $W_C = 0.5$ μm, $x = 0.3$, $N_{AB} = 1$ x 10^{18} /cm^3, $N_{DE} = 5$ x 10^{17} /cm^3, $N_{DC} = 5$ x 10^{16} /cm^3, and emitter area = 50 μm^2.

An almost identical fit is found between the numerical result and the present analytical model. A significant point to note is that the ideality factor for the curves is about 1.5 throughout. This is a result of the exponential voltage dependence of the interface transport velocity S_{EN} defined in Eqn. (5.3-6). As the base doping becomes large compared to the emitter doping, V_{JP} (and hence S_{EN}) becomes less a function of the applied voltage, and then the ideality factor approaches unity.

In Fig. 5.4 we compare the model with the experimental results of Marty et al [Mart88], on an N-$Al_xGa_{1-x}As$/p-GaAs/N-$Al_yGa_{1-y}As$ double hetrojunction transistor. Both the junctions are graded over a thickness
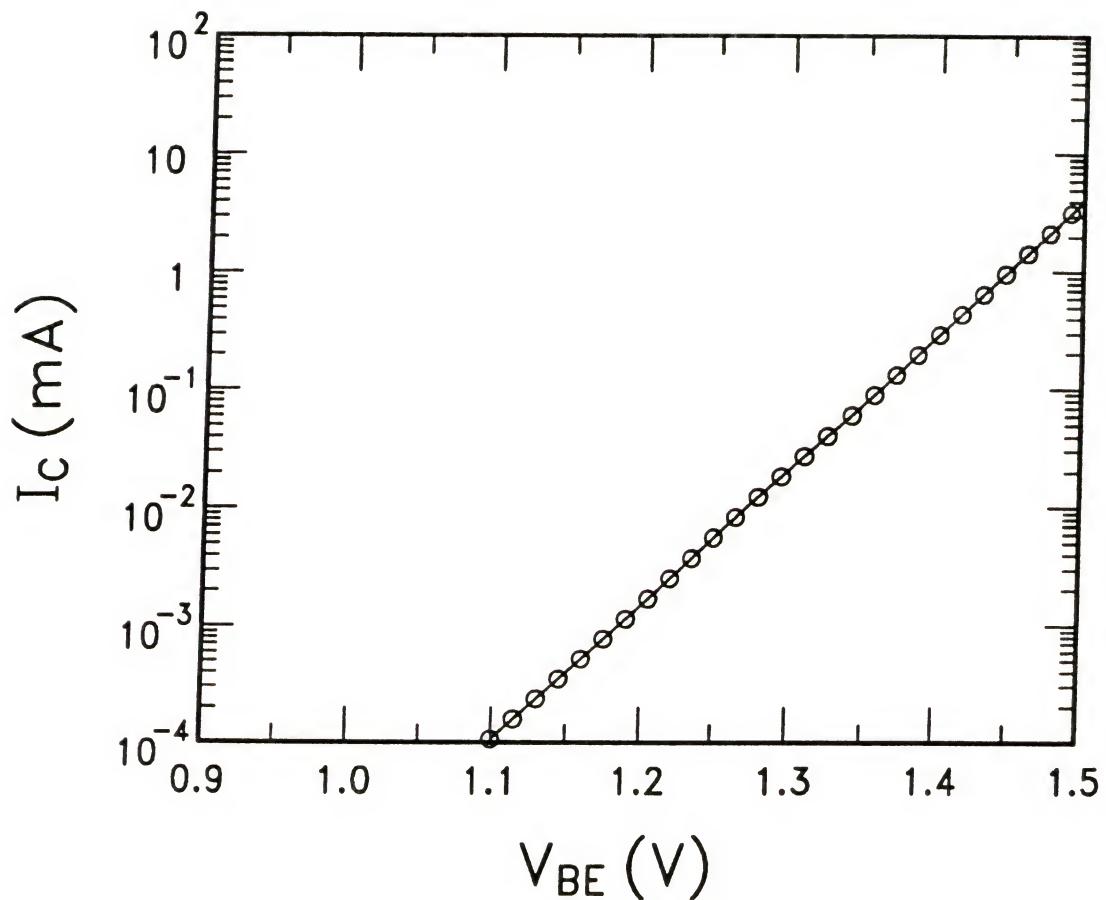
Fig. 5.3 I_C-V_{BE} characteristics from the analytical model (———)and numerical simulation results (o) [Yok84].
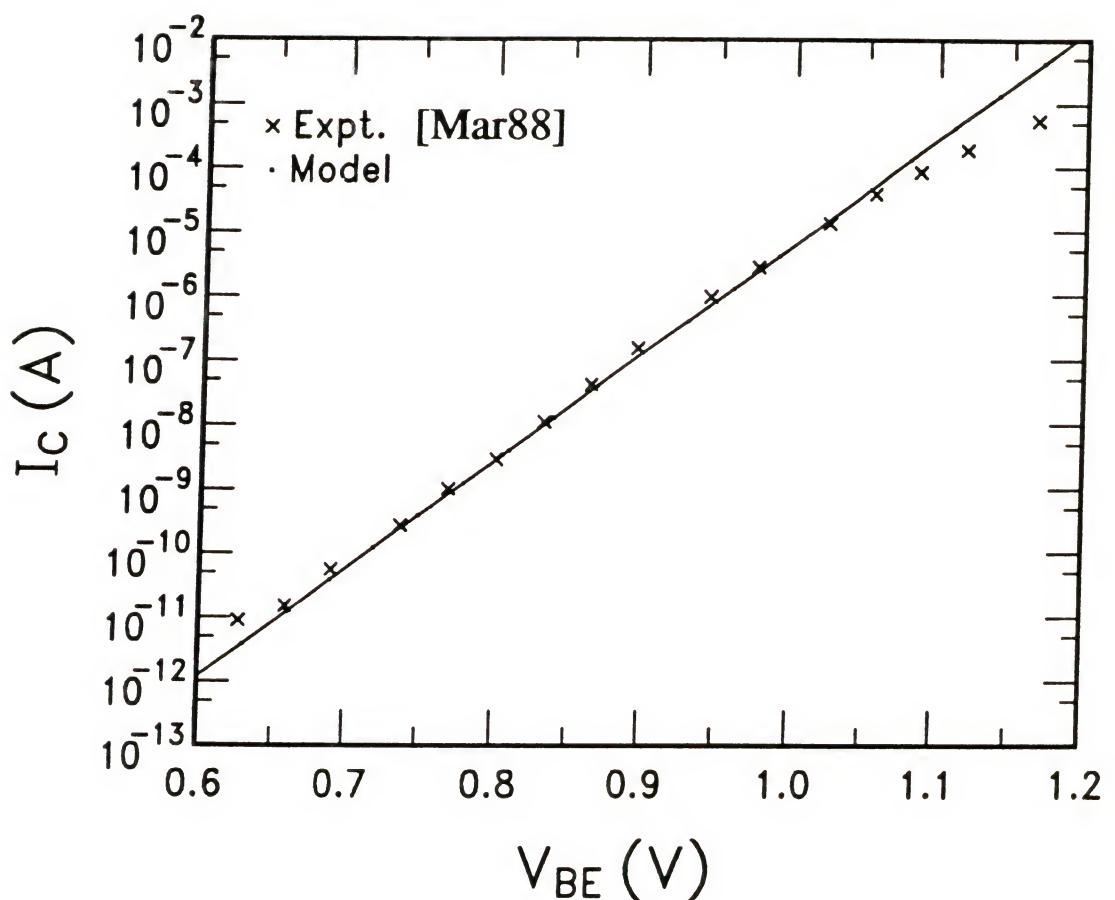
Fig. 5.4 Comparison of experimental data (x) of [Mar88] with the model.

of 800 A°. The various parameters used are as follows: $\Delta E_G = 1.55x$, $\varepsilon/\varepsilon_0$ = (13.1 - 3x), $\Delta E_C = 0.6\Delta E_G$ (as in [Wan86]), $\tau_p = \tau_n = 1$ ns (recombination lifetimes in the base), $\mu_p = 100$ cm^2/V·s, $W_E = 0.25$ μm, $W_B = 0.08$ μm, $W_C = 1.5$ μm, $x = y = 0.4$, $N_{AB} = 5$ x 10^{18} /cm^3, $N_{DE} = 10^{16}$ /cm^3, $N_{DC} = 10^{16}$ /cm^3, and emitter area = 1.2 x 10^{-4} cm^2. The mobility is used to fit the experimental data, and is found to be 2500 cm^2/V·s. An excellent fit is found between the data and the model. The ideality factor of the curves is unity. At high biases, the experimental points are seen to deviate from the ideal curve due to series resistances.

In Fig. 5.5, we compare the present model with that of Ryum and Abdel-Motaleb [Ryu90] for a transistor in which the doping density is linearly graded in the base. An abrupt N-Al$_x$Ga$_{1-x}$As/p-GaAs/n-GaAs HBT is assumed, with x = 0.3. The base doping is graded from 5 x 10^{18} /cm^3 at the emitter end to 10^{17} /cm^3 at the collector end. The rest of the parameters are the same as those for Fig. 5.4. The plots for the two models show that the model by Ryum and Abdel-Motaleb overestimates the current by a factor of almost 4.

5.6 Conclusion

In conclusion, we have derived a new charge-control expression for the collector current in a heterojunction bipolar transistor that is valid for arbitrary doping profiles in the base (and for low and moderate injection levels in an abrupt heterojunction transistor), and for both single and double heterojunction transistors. The model is shown to be an improvement over another charge-control model [Ryu90] derived recently.
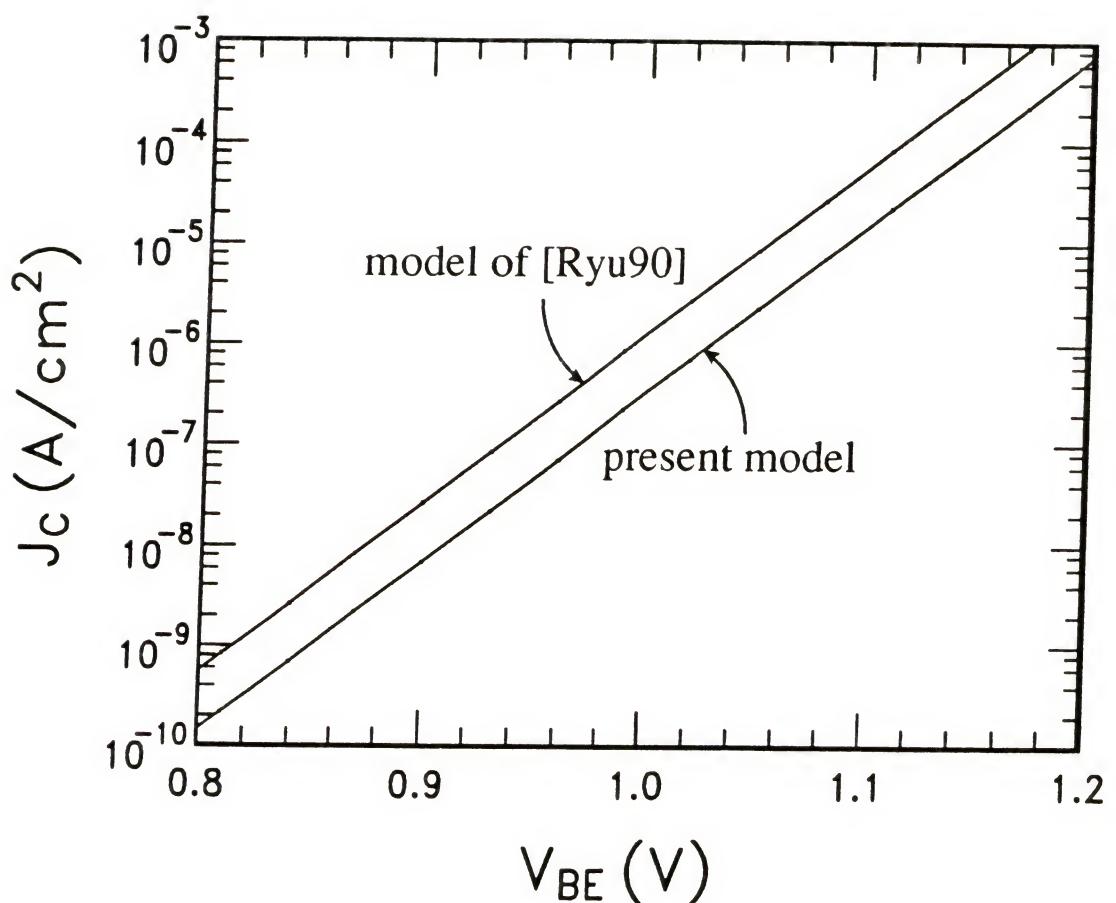
Fig. 5.5 Comparison of the present model with the model of [Ryu90] for a graded doping density in the base.

CHAPTER 6
MODELING SPACE CHARGE REGION RECOMBINATION IN HBT'S

6.1 Introduction

The purpose of this chapter is to derive a new model for the bulk space charge region (SCR) recombination current in heterojunction bipolar transistors (HBTs). There were two motivations for deriving this model:

(1) At low current densities, the base current and the current gain of silicon homojunction bipolar transistors are determined primarily by the bulk SCR recombination and the surface recombination currents [Sze81]. These currents are all the more significant in HBTs, because, in general, the compound semiconductors making up the HBT have larger bulk defect densities (and hence larger recombination-center densities) and poorer surface qualities. Moreover, many materials (particularly GaAs) used to make HBTs have larger energy gaps than the energy gap of silicon. A larger energy gap implies a smaller intrinsic carrier density, n_i (since $n_i = (N_C N_V)^{1/2} \exp(-E_G/2kT)$, where E_G is the energy gap). Now the recombination current in the quasi-neutral regions of a transistor (I_{QNR}) varies as n_i^2, while the SCR recombination current (I_{SCR}) varies as n_i, so that $I_{SCR}/I_{QNR} = 1/n_i$. Thus the smaller the n_i, the more important I_{SCR} is. Thus in HBTs that are made of GaAs, the SCR recombination current is more significant, perhaps even at high current levels, than it is in silicon. Thus modeling the SCR recombination

current is essential for predicting transistor behavior, and for developing comprehensive device and circuit simulators.

(2) The presence of a heterojunction makes the SCR recombination current models derived for the homojunction [SahC57, Shu88] inadequate. Specifically, the larger energy gap in the emitter of an HBT makes the minority carrier density much smaller than what it would be in a homojunction. This smaller density results in a smaller recombination rate in the emitter. A model for the SCR recombination current must, therefore, take into consideration the effects of the change in the energy gap from the base to the emitter. Note here that the change in the energy gap, ΔE_G, may be a function of position within the emitter, if the heterojunction is a graded-gap junction. Effects of such grading must also be accounted for in the model.

The approach we use to derive the new model is as follows:

(1) The continuity equation is integrated over the space charge region to establish mathematically that the SCR recombination current density, J_{SCR}, can be expressed as the integral of the recombination rate, $R(x)$, over the SCR.

(2) Shockley-Read-Hall statistics [Sho52, Hall52] are assumed for the recombination rate $R(x)$.

(3) The integral for R is evaluated to obtain J_{SCR}. The variations in the energy gap within the emitter are accounted for in the evaluation. This step marks the first point of distinction between the present model and the classical homojunction theory [SahC57] for SCR recombination.

The above three steps are carried out in section 6.2.

In the last chapter, we derived a new charge-control model for the HBT. In that derivation, we neglected the SCR recombination current. Thus in Sec. 6.3 below, we incorporate the model for J_{SCR} derived in Sec. 6.2, into the charge-control model of chapter 5. This process of incorporation leads to some very unique and peculiar interrelationships, which are best discussed after we have gone through the mathematics. We therefore withhold further explanation of this issue until the latter part of Sec. 6.3. These unique interrelationships mark the second point of distinction between heterojunction and homojunction bipolar transistor models.

In Sec. 6.4 we present and discuss some numerical results. The effects of the energy gap grading in the emitter on the base and the collector currents, and on the current gain are studied. The chapter is then concluded in Sec. 6.5.

6.2 Derivation of the SCR Recombination Current

The following is the list of assumptions we make in the analysis.

(1) DC steady-state conditions.

(2) One dimensional current flow.

(3) An N-$Al_yGa_{1-y}As$/p-GaAs/n-GaAs graded-gap transistor, unless otherwise mentioned. The methodology, though, is applicable to any heterojunction system.

(4) The space charge region width on the emitter side (X_{NE} in Fig. 6.1) is larger than the width (λ in Fig. 6.1) over which the energy gap is graded. This is the most general case possible from a modeling point of view. The case when $X_{NE} < \lambda$, is a special case of the above.
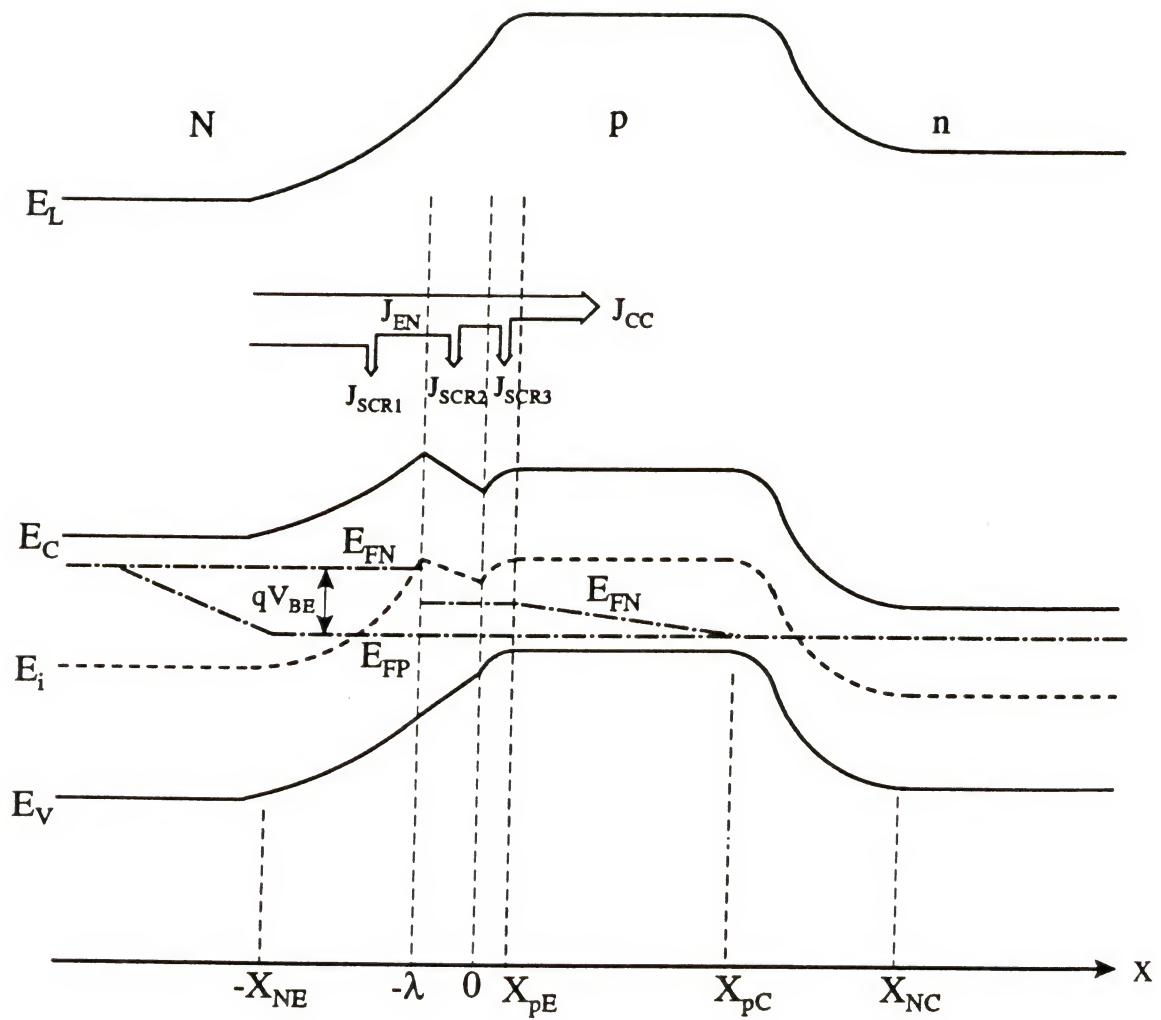
Fig. 6.1 Energy band diagram of an N-AlGaAs/p-GaAs/n-GaAs graded gap heterojunction bipolar transistor (not drawn to scale). The figure also illustrates the effect of SCR recombination on the electron current; note that $J_{EN} = J_{SCR2} + J_{SCR3} + J_{CC}$.

(5) The hole quasi-Fermi level E_{FP} is nearly constant within the SCR. The electron quasi-Fermi level, E_{FN}, is nearly constant except for a possible discontinuity at the point within the SCR where there is a peak (at $x = -\lambda$ in Fig. 6.1) in the conduction band edge. This is the same as assumption (4) of chapter 5.

(6) The recombination rate $R(x)$ within the space charge region can be described by Shockley-Read-Hall statistics [Sho52, Hal152]. The presence of a peak in the conduction band may cause electrons to become hot as they pass over the spike. This may cause an Auger recombination current. Such recombination mechanisms are neglected.

(7) The hole and electron recombination lifetimes (τ_{p0} and τ_{n0}), and the energy level (E_T) of the recombination centers are not functions of position within the SCR.

(8) The intrinsic carrier density, n_i, is given by $n_i = (N_C N_V)^{1/2} \exp(-E_G/2kT)$. In the region of the emitter where the AlGaAs composition is graded, E_G, N_C and N_V will be functions of x. But the x-dependence of E_G affects n_i much more strongly than does the x-dependence of N_C and N_V. Hence the x-dependence of N_C and N_V is assumed negligible.

Fig. 6.1 shows the energy band diagram of an N-$Al_y Ga_{1-y} As$/p-GaAs/n-GaAs HBT. The energy gap is graded over a width λ in the emitter, from $x = 0$ to $x = -\lambda$ (that is, the Al fraction, y, is a function of x such that $y = 0$ at $x = 0$, and $y = y_{max}$ at $x = -\lambda$. For $x < -\lambda$, y remains constant at y_{max}. The fraction y is usually a linear or parabolic function of x in the graded region [Mey88]). The local vacuum level, E_L, represents the variation of the electrostatic potential within the

device (as we discussed in chapter 2; see also [Marsh78a]). Note that the conduction and the valence bands, E_C and E_V, are parallel to E_L everywhere except between $x = -\lambda$ and $x = 0$, where the grading in the energy gap exists [Marsh78a]. Along the x-axis, $x=0$ defines the metallurgical emitter-base interface. X_{NE} and X_{pE} are the edges of the SCR on the N (emitter) and p (base) side of the emitter-base junction. X_{NC} and X_{pC} are similar quantities at the collector-base junction. The emitter-base junction is shown forward biased, and the collector-base junction short-circuited. E_{FP} and E_{FN} are the hole and electron quasi-Fermi levels.

Now we derive a model for J_{SCR} by carrying out the three steps we outlined in Sec. 6.1.

Step 1. In steady-state, the continuity equation for electrons is written as

$$\frac{1}{q} \frac{dJ_N(x)}{dx} - R = 0$$

where R is the recombination rate. Integrating over the emitter-base SCR yields

$$\int_{-X_{NE}}^{X_{pE}} dJ_N = q \int_{-X_{NE}}^{X_{pE}} R \, dx$$

Hence

$$J_N(X_{pE}) - J_N(-X_{NE}) = q \int_{-X_{NE}}^{X_{pE}} R \, dx$$

Now

$$J_N(-X_{NE}) = J_{SCR} + J_N(X_{pE})$$

since electrons flowing from the emitter into the SCR either recombine within the SCR or exit the SCR into the base. Combining the last two equations, we get the expression for J_{SCR}:

$$J_{SCR} = - q \int_{-X_{NE}}^{X_{pE}} R \, dx \qquad (6.2-1)$$

The same equation can also be derived from the hole continuity equation.

Step 2. The recombination rate R is written assuming the classical Shockley-Read-Hall statistics [Sho52]:

$$R = \frac{pn - n_i^2}{\tau_{po} (n+n_1) + \tau_{no} (p+p_1)}$$

where p and n are the hole and electron densities respectively ($p = n_i \exp[(E_i-E_{FP})/kT]$, $n = \exp[(E_{FN}-E_i)/kT]$), n_i is the intrinsic carrier density, τ_{no} and τ_{po} are the electron and hole lifetimes within the SCR respectively, $p_1 = n_i \exp[(E_i-E_T)/kT]$, $n_1 = n_i \exp[(E_T-E_i)/kT]$ where E_T is the energy level of the recombination center (we assume a single, discrete level), and E_i is the intrinsic Fermi level. R can alternately be written as [SahC57]

$$R = \frac{n_i}{\sqrt{\tau_{po}\tau_{no}}} \frac{\sinh\left(\frac{E_{FN}-E_{FP}}{2kT}\right)}{\left\{ \cosh\left[\frac{1}{kT}\left(E_i - \frac{E_{FN}+E_{FP}}{2}\right) + \ln\sqrt{\tau_{no}/\tau_{po}}\right] + \right.}$$

$$\left. + \exp\left[(E_{FN}-E_{FP})/2kT\right] \cosh\left[\frac{E_T-E_i}{kT} + \ln\sqrt{\tau_{po}/\tau_{no}}\right] \right\} \qquad (6.2-2)$$

Step 3. We now split the integral of (6.2-1) into three parts:

$$J_{SCR} = q \int_{-X_{NE}}^{-\lambda} R \, dx + q \int_{-\lambda}^{0} R \, dx + q \int_{0}^{X_{pE}} R \, dx = J_{SCR1} + J_{SCR2} + J_{SCR3} \qquad (6.2-3)$$

In conventional homojunction theory [SahC57], the integral of Eqn. (6.2-1) would be split into two parts corresponding to the n and p sides of the SCR. In a graded-gap heterojunction, one more partition becomes necessary, to separate the graded energy gap and the constant energy gap regions in the emitter.

Now to evaluate the integrals in Eqn. (6.2-3), we need to know the position dependence of all the quantities appearing in the expression for R in Eqn. (2). Of these quantities, assumptions (5) and (7) enable us to treat E_{FN}, E_{FP}, τ_{n0}, τ_{p0} and E_T as constants within each of these integrals. We note once again that E_{FN} may have a discontinuity at $x = -\lambda$ if the current flow across that plane is by thermionic emission (assumption (5)); but it can still be assumed constant within each of the three regions of integration. Thus only two quantities, n_i and E_i, are left in the expression for R whose x-dependence must be determined. To determine this dependence, we first need to define the position-dependence of the energy gap. In an $Al_yGa_{1-y}As/GaAs$ junction, if E_{GB} is the energy gap of GaAs, then the energy gap of $Al_yGa_{1-y}As$, E_{Gy}, is given by [Yos85]

$$E_{Gy} \simeq E_{GB} + 1.55y \qquad\qquad (6.2\text{-}4a)$$

Thus

$$\Delta E_G \stackrel{\Delta}{=} E_{Gy} - E_{GB} \simeq 1.55y \qquad\qquad (6.2\text{-}4b)$$

Now as we mentioned earlier in this section, the functional dependence of $y(x)$ is such that

$$y(x>0) = 0 \text{ and } y(x<-\lambda) = y_{max}$$

Thus the position dependence of $E_G(x)$ can be written as

$$E_G = E_{GB} \qquad\qquad x \geq 0 \qquad\qquad (6.2\text{-}5a)$$

$$= E_{GB} + 1.55\, y(x) \qquad -\lambda \leq x \leq 0 \qquad (6.2\text{-}5b)$$

$$= E_{GB} + 1.55\, y_{max} \qquad x \leq -\lambda \qquad (6.2\text{-}5c)$$

Thus E_G is a function of x only in the region $-\lambda \leq x \leq 0$. Now n_i is given by

$$n_i = (N_C N_V)^{1/2}\, \exp(-E_G/2kT)$$

From this and assumption (8), we can write down the position dependence of n_i:

$$n_i = n_{iB} = (N_{CB} N_{VB})^{1/2})\, \exp(-E_{GB}/2kT) \qquad x \geq 0 \qquad (6.2\text{-}6a)$$

$$= n_{iB}\, \exp(-\Delta E_G/2kT)$$

$$= n_{iB}\, \exp\,[-1.55 y(x)/2kT] \qquad -\lambda \leq x \leq 0 \quad (6.2\text{-}6b)$$

$$= n_{iE} = n_{iB}\, \exp\,(-1.55 y_{max}/2kT) \qquad x \leq -\lambda \quad (6.2\text{-}6c)$$

where the subscript B refers to the base region. Thus n_i is a function of x only in the second integral of Eqn. (6.2-3).

Now we determine E_i - the intrinsic Fermi level - as a function of x. It has been shown [Marsh78a, Lund82] that E_i in a semiconductor of nonuniform composition can be written as

$$E_i(x) = E_0 - \chi(x) - E_G(x)/2 + kT/2\, \ln\,[N_V(x)/N_C(x)] - qV(x) \quad (6.2\text{-}7)$$

where E_0 is a reference energy level, χ is the electron affinity, and V is the electrostatic potential (note that Eqn. (6.2-7) assumes the validity of the electron affinity rule that we discussed in chapter 2). Also, the conduction band edge $E_C(x)$ is given by [Lund82]

$$E_C(x) = E_0 - \chi(x) - qV(x) \qquad\qquad (6.2\text{-}8)$$

We now define the zero references of $E_i(x)$ and $V(x)$ as the E_i and V at x = X_{pE}, respectively. Thus

$$V \ (x=X_{pE}) \ \overset{\Delta}{=} \ 0$$

$$E_i(x=X_{pE}) \ \overset{\Delta}{=} \ 0 \qquad\qquad\qquad (6.2\text{-}9)$$

so that

$$E_0 = \chi(X_{pE}) + E_G(X_{pE})/2 + kT/2 \ \ln \ [N_V(X_{pE})/N_C(X_{pE})] \qquad (6.2\text{-}10)$$

Thus $E_i(x)$ can be rewritten as

$$E_i(x) = -\Delta\chi(x) - \Delta E_G(x)/2 - qV(x) \qquad\qquad (6.2\text{-}11)$$

where $\Delta\chi(x) = \chi(x) - \chi(X_{pE})$, and as stated in assumption (8), we have

neglected the variations in N_C and N_V. Now it has also been shown

[Wan86] that in an AlGaAs/GaAs system,

$$\Delta\chi = - \ 0.6 \ \Delta E_G \qquad\qquad\qquad (6.2\text{-}12)$$

Therefore

$$E_i(x) = 0.1 \ \Delta E_G(x) - qV(x) \qquad\qquad (6.2\text{-}13)$$

Now normalizing with kT, and substituting Eqns. (6.2-4) and (6.2-5),

yields

$$E_i(x)/kT = -q/kT \ V(x) \ \overset{\Delta}{=} \ -u(x) \qquad\qquad x \geq 0 \qquad (6.2\text{-}14a)$$

$$= 0.155 \ y(x) - u(x) \qquad -\lambda \leq x \leq 0 \quad (6.2\text{-}14b)$$

$$= 0.155 \ y_{max} - u(x) \qquad x \leq -\lambda \qquad (6.2\text{-}14c)$$

where $u(x)$ is the non-dimensional electrostatic potential. Eqns. (6.2-6)

and (6.2-14) respectively define the x-dependence of n_i and E_i, as we

had desired.

We now rewrite Eqn. (6.2-2) as

$$R = \frac{\sinh[(E_{FN}-E_{FP})/2kT]}{\tau} \ \frac{n_i(x)}{\cosh \ [E_i(x)/kT-U_F] + b} \qquad (6.2\text{-}15)$$

where

$$U_F = (E_{FN}+E_{FP})/2kT - 1/2 \ \ln(\tau_{no}/\tau_{po}) \qquad\qquad (6.2\text{-}15a)$$

$$b = \exp[(E_{FP}-E_{FN})/kT] \cosh [(E_T-E_i)/kT + 1/2 \ln(\tau_{po}/\tau_{no})] \quad (6.2\text{-}15b)$$

$$\text{and} \quad \tau = (\tau_{po}\tau_{no})^{1/2} \quad (6.2\text{-}15c)$$

We now return to the integrals of Eqn. (6.2-3). Of the three integrals in that equation, only in the second one (for the range $-\lambda < x < 0$) are both n_i and E_i dependent on x; in the first and third integrals, only E_i is a function of x. We found the second integral not amenable to analytical evaluation, and thus left it to be evaluated numerically. From Eqns. (6.2-6b,14b,15) J_{SCR2} of Eqn. (6.2-3) is written as

$$J_{SCR2} = \frac{\sinh [(E_{FN}-E_{FP})/kT]}{\tau} n_{iB} \int_{-\lambda}^{0} \frac{\exp [1.55y(x)/2kT]}{\cosh [0.155y(x)-u(x)-U_F] + b} dx \quad (6.2\text{-}16)$$

Let us now examine the integral for J_{SCR1} of Eqn. (6.2-3). From Eqns. (6.2-3,6c,14c,15)

$$J_{SCR1} = \frac{\sinh [(E_{FN}-E_{FP})/kT]}{\tau} n_{iE} \int_{-X_{NE}}^{-\lambda} \frac{dx}{\cosh [u(x)-0.155y_{max}+U_F] + b} \quad (6.2\text{-}17)$$

where we have used the identity $\cosh(-z) = \cosh(z)$. To evaluate the integral, we must know u(x) as a function of x. For this we make an approximation. In the denominator of the integrand, if b is small compared to the cosh term (which it will be if $E_T \simeq E_i$), then the integral will have a local maximum in the interval $[-X_{NE},-\lambda]$ and will fall off exponentially away from that point; thus the major contribution to the integral will come from the region near the maximum. We therefore

expand u(x) in a Taylor series about the maximum and retain only the linear term:

$$u(x) \simeq u(X_{01}) + (x-X_{01}) \, du(X_{01})/dx \qquad (6.2\text{-}18)$$

where X_{01} is the point at which the integrand has a maximum in the interval $[-X_{NE},-\lambda]$, and $du(X_{01})/dx$ is $du(x)/dx$ evaluated at $x=X_{01}$. Eqn. (6.2-17) can now be written as

$$J_{SCR1} = \frac{n_{iE} \, \sinh[(E_{FN}-E_{FP})/kT]}{\tau} \int_{-X_{NE}}^{-\lambda} \frac{dx}{\{ \cosh[(x-X_{01})du/dx + u(X_{01}) - 0.155y_{max} + U_F] + b \}} \qquad (6.2\text{-}19)$$

This integral can now be evaluated analytically. The final result is

$$J_{SCR1} = \frac{n_{iE} \, \sinh[qV_{BE}/kT]}{\tau} \frac{2}{(1-b_1{}^2)^{1/2}} \frac{1}{du/dx}$$

$$\left\{ \tan^{-1}\left[\sqrt{(1-b_1)\Big/(1+b_1)} \ \tanh\left(\frac{u(-\lambda)-\alpha\lambda+U_F}{2} \right) \right] - \right.$$

$$\left. \tan^{-1}\left[\sqrt{(1-b_1)\Big/(1+b_1)} \ \tanh\left(\frac{u(-X_{NE})-\alpha\lambda+U_F}{2} \right) \right] \right\} \quad ..\text{if } b_1{<}1$$

$$\ldots(6.2\text{-}20)$$

where we have used $(E_{FN}-E_{FP}) = qV_{BE}$ (for $-X_{NE} \leq x \leq -\lambda$; see Fig. 6.1). J_{SCR3} can be derived similarly:

$$J_{SCR3} = \frac{n_{iB} \, \sinh[(E_{FN}-E_{FP})/2kT]}{\tau} \frac{2}{(1-b_3{}^2)^{1/2}} \frac{1}{du(X_{03})/dx}$$

$$\left\{ \tan^{-1}\left[\sqrt{(1-b_3)\Big/(1+b_3)} \ \tanh\left(\frac{u(X_{pE})+U_F}{2} \right) \right] - \right.$$

$$\left. \tan^{-1}\left[\sqrt{(1-b_3)\Big/(1+b_3)} \ \tanh\left(\frac{u(0)+U_F}{2} \right) \right] \right\} \quad ..\text{if } b_3{<}1$$

$$(6.2\text{-}21)$$

In Eqns. (6.2-20) and (6.2-21), $u(x)$ and du/dx at the various points are evaluated by solving Poisson's equation using the depletion approximation [And62]. X_{01} is the value of x at which $[u(x)-\alpha\lambda+U_F]$ has a minimum in the closed interval $[-X_{NE},-\lambda]$; it is thus evaluated by solving the equation, $d[u(x)-0.155y_{max}+U_F]/dx = 0$. X_{03} is found similarly. Equations (6.2-3,16,20,21), when put together, constitute the final expression for J_{SCR}. We note once again that for J_{SCR2} and J_{SCR3}, $(E_{FN}-E_{FP})$ is less than qV_{BE} (assumption (5), Fig. 6.1). We will determine it in the next section. For future reference, we utilize the relation

$$p(X_{pE})n(X_{pE}) = n_{iB}^2 \exp[(E_{FN}-E_{FP})/kT]$$

to write

$$J_{SCR2} + J_{SCR3} = k_R [p(X_{pE})n(X_{pE})]^{1/2} \qquad (6.2-22)$$

where

$$k_R = \frac{q}{2\tau} \int_{-\lambda}^{0} \frac{dx}{\cosh[-u(x)-\alpha x-U_F] + b} + \frac{q}{\tau (1-b_3^2)^{1/2} du(X_{03})/dx}$$

$$\left\{ \tan^{-1}\left[\sqrt{(1-b_3)\Big/(1+b_3)} \ \tanh\left(\frac{u(X_{pE})+U_F}{2} \right) \right] - \right.$$

$$\left. \tan^{-1}\left[\sqrt{(1-b_3)\Big/(1+b_3)} \ \tanh\left(\frac{u(0)+U_F}{2} \right) \right] \right\} \qquad (6.2-22a)$$

6.3 Incorporation of J_{SCR} Into the Charge-Control Model for the HBT

In this section we derive a modified charge control model that includes the SCR recombination current, and will discover and discuss the peculiarities of this incorporation for an HBT.

The electron current density (J_{CC}) flowing through the quasi-neutral base of any bipolar transistor can be written as (see Eqn. (5.3-3b))

$$J_{CC} = q\ D_n\ \frac{(pn)_{x=X_{pC}} - (pn)_{x=X_{pE}}}{\int_{X_{pE}}^{X_{pC}} p\ dx} \tag{6.3-1}$$

In deriving Eqn. (6.3-1), recombination in the quasi-neutral base (QNB) is neglected, and the energy gap is assumed constant (in the QNB). The pn product at a heterojunction, as we saw in chapter 5 (see Eqn. (5.3-11)), satisfies the following relation:

$$p(X_{pE})n(X_{pE}) - (J_{EN}/qS_{EN})\ p(X_{pE}) = n_{iB}{}^2\ exp(qV_{BE}/kT) \tag{6.3-2}$$

where $J_{EN} = J_N(x = -\lambda)$ is the electron current flowing across the conduction band peak ($x = -\lambda$, see Fig. 6.1). At the base-collector junction we can similarly write

$$p(X_{pC})n(X_{pC}) + (J_{CN}/qS_{CN})\ p(X_{pC}) = n_{iB}{}^2\ exp(qV_{BC}/kT) \tag{6.3-3}$$

if that junction is also a heterojunction, which we assume for this section. Now if there were no SCR recombination current present (as we assumed in chapter 5 - assumption (5) in Sec. 5.2), J_{EN} of Eqn. (6.3-2) would be equal to J_{CC} of Eqn. (6.3-1). But in the presence of a J_{SCR} we must write (Fig. 6.1)

$$J_{EN} = J_{SCR2} + J_{SCR3} + J_{CC} \tag{6.3-4}$$

Now equations (6.2-22,6.3-1,2,3,4) must be solved simultaneously to determine the final solution for J_{CC} and J_{SCR}. These equations can be combined to yield a quadratic equation either for J_{CC} or for the quantity $[p(X_{pE})n(X_{pE})]^{1/2}$, which can subsequently be solved. The

mathematics involved in such a derivation is cumbersome but straightforward, and hence is not included here. The final expression for J_{CC} is obtained as

$$J_{CC} = -\frac{q^2 D_n n_{iB}^2 \left[\exp(qV_{BE}/kT) - \exp(qV_{BC}/kT) \right]}{Q_B} \left[-\frac{d^2}{2q^2 D_n c} \frac{Q_C}{Q_B} + 1 - \sqrt{\frac{d^2 n_{iVE}^2 Q_C}{c^2 Q_B} \left(1 + \frac{d^2 Q_C}{4q^4 D_n^2 n_{iVE}^2 Q_B} \right)} \right]$$

$$(6.3\text{-}5)$$

where

$$n_{iVE} = n_{iB} \exp(qV_{BE}/2kT) \qquad n_{iVC} = n_{iB} \exp(qV_{BC}/2kT) \qquad (6.3\text{-}6a)$$

$$Q_B = q \int_{X_{pE}}^{X_{pC}} p \; dx + \frac{qD_n}{S_{EN}} N_{AB}(X_{pE}) + \frac{qD_n}{S_{CN}} N_{AB}(X_{pC}) \qquad (6.3\text{-}6b)$$

$$Q_C = q \int_{X_{pE}}^{X_{pC}} p \; dx + \frac{qD_n}{S_{CN}} N_{AB}(X_{pC}) \qquad (6.3\text{-}6c)$$

$$c = q^2 D_n (n_{iVE}^2 - n_{iVC}^2) \qquad d = (qD_n/S_{EN}) N_{AB} X_{pE}) k_R \qquad (6.3\text{-}6d)$$

where $N_{AB}(x)$ is the doping density in the base. On the right hand side of Eqn. (6.3-5) the term in front of the square brackets is the electron current if there were no SCR recombination (see Eqn. (5.3-13)). The term inside the brackets represents the effect of the SCR recombination current on the quasi-neutral base transport. In a homojunction transistor, the SCR recombination current has no such effect on J_{CC}. In particular, the factor k_R, as defined in Eqns. (6.2-22), would have had no effect on J_{CC} in a homojunction transistor. While in a heterojunction transistor, k_R appears in the expression for J_{CC} in Eqn. (6.3-5) through Eqn. (6.3-6d). This is thus unique to the heterojunction transistor.

A brief qualitative explanation of this interrelationship is as follows: In an HBT, if there is a large peak in the conduction band, the electron flow is hindered across the energy barrier created due to this peak. Current flow across the point of the peak then, becomes the most significant bottleneck in determining the transistor characteristics. (In an n/p/n homojunction transistor, the bottleneck is the electron flow in the quasi-neutral base region). Thus for a given bias V_{BE}, the current is fixed by the flow over the peak ($x = -\lambda$ in Fig. 6.1). Then if the SCR recombination current in the region $x > -\lambda$ is increased (by decreasing the lifetime, for example), then the current flowing through the quasi-neutral base will decrease, because the sum of the two must remain a constant. This is how the SCR recombination current affects the QNB current. We note two points though: (1) If the emitter-base junction is graded sufficiently, then no peak would exist; then current flow through the SCR would be as in a homojunction, and no interrelationship will exist. The term within the square brackets in Eqn. (27) will then be unity. (2) The peak height can vary from zero to many thermal voltages; and in general, the transport through the base would be as important a factor as the current transport over the peak in determining the device characteristics. Thus the above explanation represents just one extreme case.

Now the term inside the brackets in Eqn. (6.3-5), as it stands, is a function of the injection level in the base. Such a dependence makes the explicit determination of J_{CC} intractable. Therefore, realizing that in practical HBT's the base doping is larger than the

emitter doping, the base is unlikely to go into high injection. Hence we assume that all the quantities inside the brackets in Eqn. (6.3-5) can be evaluated under low injection conditions in the base. Then, Q_B, Q_C and d in Eqns. (6.3-6) can be written as

$$Q_B = qN_{AB}W_B'[1+(D_n/W_B')(1/S_{EN}+1/S_{CN})] \qquad (6.3\text{-}7a)$$

$$Q_C = qN_{AB}W_B' (1 + D_n/W_B'S_{CN}) \qquad (6.3\text{-}7b)$$

$$d = qD_nN_{AB}k_R/S_{EN} \qquad (6.3\text{-}7c)$$

where $W_B' = (X_{pE}-X_{pC})$. Eqn. (6.3-5) can now be rewritten as

$$J_N = - \frac{q^2D_n(n_{iVE}{}^2-n_{iVC}{}^2)}{Q_B} \left[1 + \frac{RT^2r_s}{2r_n} \sqrt{\frac{RT^2r_s}{r_n{}^2} \left(1 + \frac{RT^2r_s}{4} \right)} \right] \qquad (6.3\text{-}8)$$

where

$$RT = \frac{k_RN_{AB}}{qS_{EN}n_{iVE}} \qquad r_s = \frac{Q_C}{Q_B} \qquad \text{and} \qquad r_n = 1 - \frac{n_{iVC}{}^2}{n_{iVE}{}^2} \qquad (6.3\text{-}9)$$

To determine the SCR recombination currents of Eqns. (6.2-20,21,22), we need the pn product at $x = X_{pE}$. This is found to be

$$[p(X_{pE})n(X_{pE})]^{1/2} = n_{iB} \exp(qV_{BE}/2kT) \, r_s{}^{1/2} \, [-RT \, r_s{}^{1/2} + (RT^2r_s+1)^{1/2}]$$

$$(6.3\text{-}10)$$

after assuming low injection for quantities inside the square brackets.

This completes the modeling aspect of this chapter.

6.4 Numerical Results

In this section we present some numerical results obtained using the model derived above. The approach here is more of investigation rather than prediction of HBT characteristics, since little is known about the lifetime behaviors within the SCR. For the results below, the

electron and hole lifetimes are assumed to be equal and constant throughout the entire SCR.

The following parameter values are used for the numerical results shown in Figs. 6.2 and 6.3. The transistor is an $N-Al_yGa_{1-y}As/p-GaAs/n-GaAs$ type. The various parameters used are: $\Delta E_G = 1.55y$, $\varepsilon/\varepsilon_o = (13.1 - 3y)$, $\Delta E_C = 0.6\Delta E_G$, $\tau_p = \tau_n = 1$ ns (recombination lifetimes), $\mu_n = 3090$ $cm^2/V \cdot s$, $\mu_p = 200$ $cm^2/V \cdot s$, $W_E = 0.2$ μm, $W_B = 0.1$ μm, $W_C = 0.2$ μm, $y = 0.3$, $N_{AB} = 5 \times 10^{18}$ $/cm^3$, $N_{DE} = 10^{17}$ $/cm^3$, and $N_{DC} = 10^{17}$ $/cm^3$.

Fig. 6.2 shows the SCR recombination current as a function of the base-emitter voltage (V_{BE}) for three different grading widths (λ). As the grading width is increased, the recombination current is seen to increase. This is because a graded energy gap allows greater hole injection into the emitter SCR than does an abrupt one, so that there is larger recombination in the emitter SCR (specifically, the J_{SCR2} term in Eqn. (6.2-3) increases, as we will see below).

Figs. 6.3(a), 6.3(b) and 6.3(c) each show the components of J_{SCR} (as defined by Eqn. (6.2-3)) alongwith J_{SCR} itself, for grading widths of 0 A°, 100 A° and 300 A° respectively. Tables 6.1(a), 6.1(b) and 6.1(c) show the corresponding ideality factors for the current components at three different base-emitter voltages.

Fig. 6.3(a) shows the components for an abrupt base-emitter junction. The component J_{SCR2} is not shown since for an abrupt junction it is zero. We see from the figure that J_{SCR1} dominates over J_{SCR3} for most of the voltage range, although at higher voltages the latter catches up since it has a smaller ideality factor. The dominance of
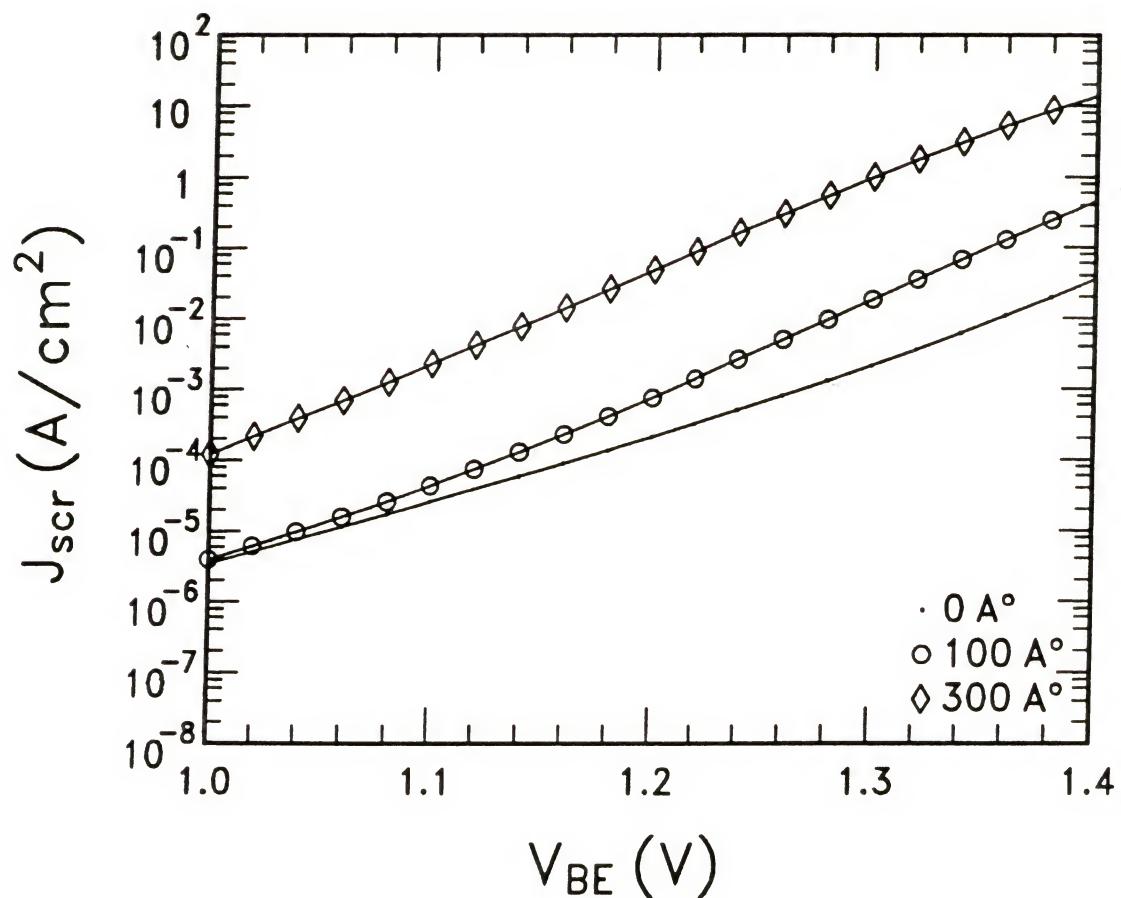
Fig. 6.2 SCR recombination current versus the base-emitter voltage for three different grading thicknesses in the emitter.
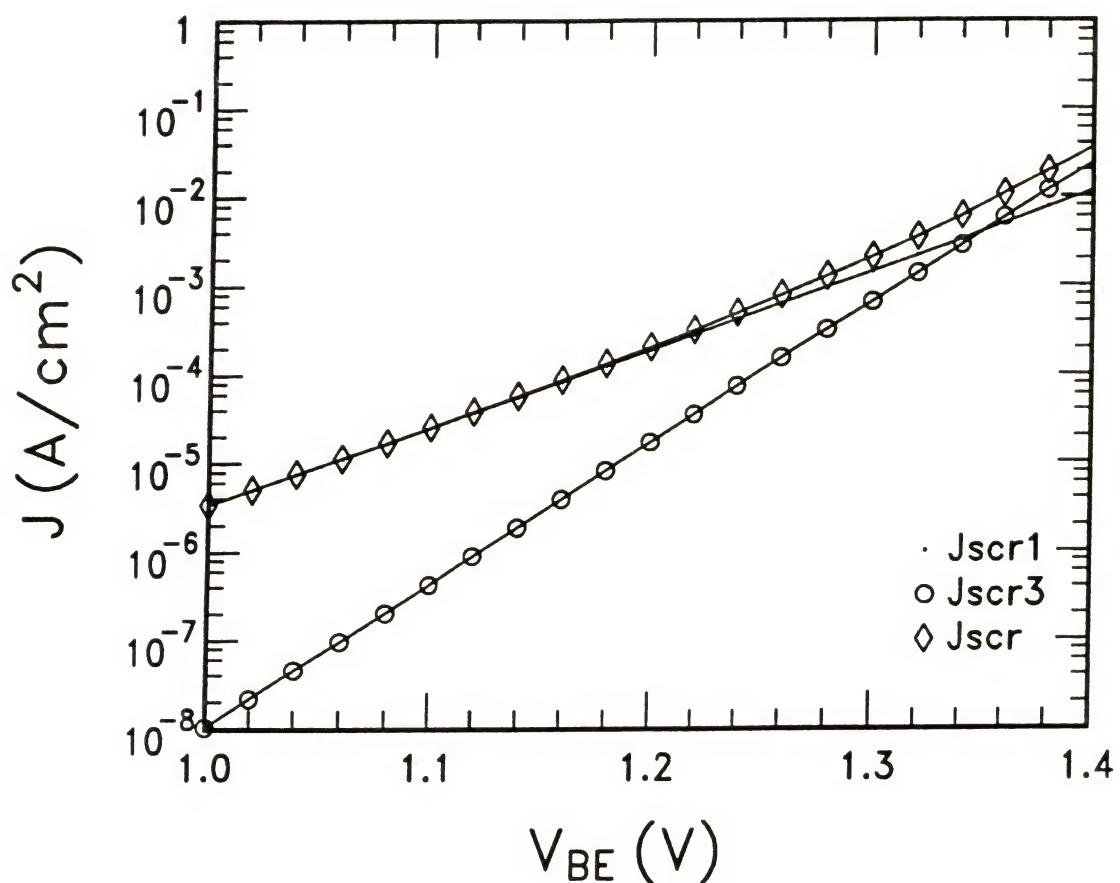
Fig. 6.3(a) Components of the SCR recombination current vs. the base-emitter voltage for an emitter graded layer thickness (λ) of 0 A°.

$V_{BE} =$	1 V	1.2 V	1.4 V
J_{SCR1}	1.93	1.91	1.83
J_{SCR2}	--	--	--
J_{SCR3}	1.04	1.05	1.11
J_{SCR}	1.93	1.80	1.29

Table 6.1(a) Ideality factors of SCR recombination current components for different base-emitter voltages, for a graded layer thicknesses of 0 A°.

J_{SCR1} (which is the recombination on the emitter side) is due to the large base to emitter doping density ratio, which causes most of the SCR to lie on the emitter side. In general, there are two competing factors affecting the SCR recombination in an abrupt heterojunction. One is the ratio of the doping densities on the base and the emitter sides, and the other is the difference in the energy gaps of the base and the emitter. For the particular choice of parameters used for Fig. 6.3(a), the former is seen to dominate, so that the emitter SCR recombination dominates over the base SCR recombination; if the latter were to dominate in a particular heterojunction, the base SCR recombination will be larger.

Fig. 6.3(b) shows the recombination current components for a grading width of 100 A°. At lower biases, J_{SCR1} is seen to dominate the recombination current; as the bias increases, J_{SCR2} becomes dominant. This is primarily because with increasing forward bias, the interval [-X_{NE},λ] becomes smaller and thus J_{SCR1} does not increase as rapidly as J_{SCR2} (see Fig. 6.1 and Eqn. (6.2-3)). This decreasing X_{NE} also explains the large ideality factor for J_{SCR1} in Table 6.1.

Fig. 6.3(c) shows the J_{SCR} components for a grading width of 300 A°. At this width, the recombination current is entirely determined by the recombination in the graded emitter SCR. But we note that such behavior occurs only when the base doping density is much higher than the emitter doping density, as we discussed above. In Fig. 6.4 we show the variation of recombination current density in the graded emitter region (J_{SCR2}) and the total recombination current density (J_{SCR}) as a function of the graded layer width (λ). The base-emitter bias is kept
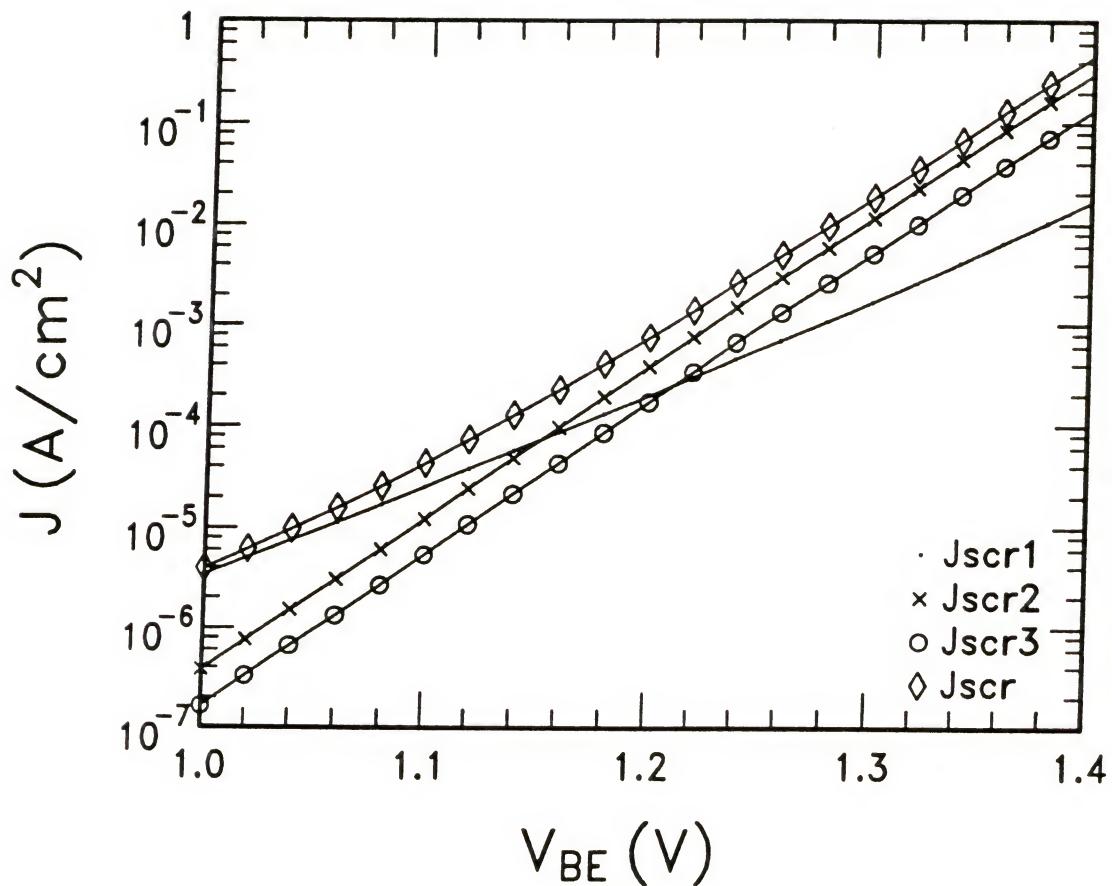
Fig. 6.3(b) Components of the SCR recombination current vs. the base-
emitter voltage for an emitter graded layer thickness (λ) of
100 A°

$V_{BE} =$	1 V	1.2 V	1.4 V
J_{SCR1}	1.93	1.87	1.65
J_{SCR2}	1.11	1.13	1.24
J_{SCR3}	1.10	1.13	1.23
J_{SCR}	1.73	1.27	1.24

Table 6.1(b) Ideality factors of SCR recombination current components for different base-emitter voltages, for a graded layer thicknesses of 100 A°.
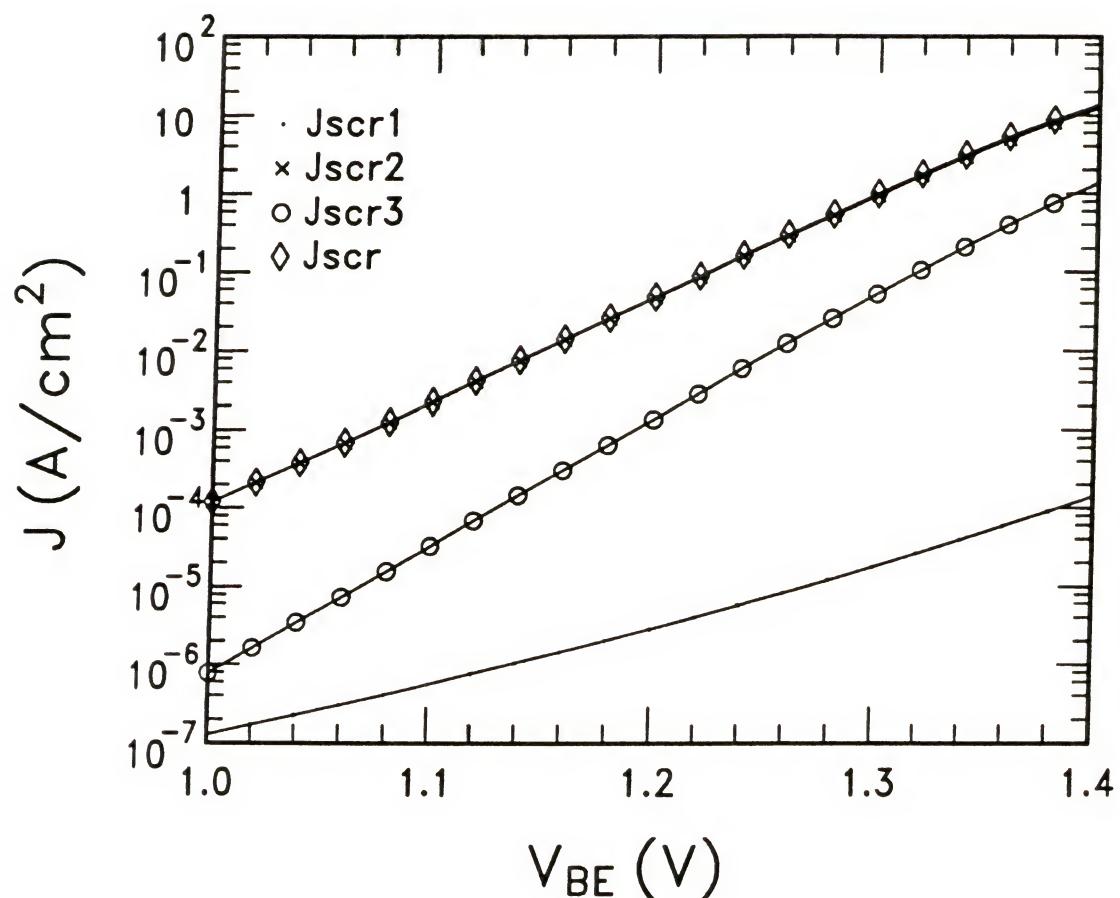
Fig. 6.3(c) Components of the SCR recombination current vs. the base-emitter voltage for an emitter graded layer thickness (λ) of 300 A°.

$V_{BE} =$	1 V	1.2 V	1.4 V
J_{SCR1}	2.72	2.26	1.75
J_{SCR2}	1.33	1.27	1.80
J_{SCR3}	1.03	1.04	1.31
J_{SCR}	1.33	1.26	1.74

Table 6.1(c) Ideality factors of SCR recombination current components for different base-emitter voltages, for a graded layer thicknesses of 300 A°.
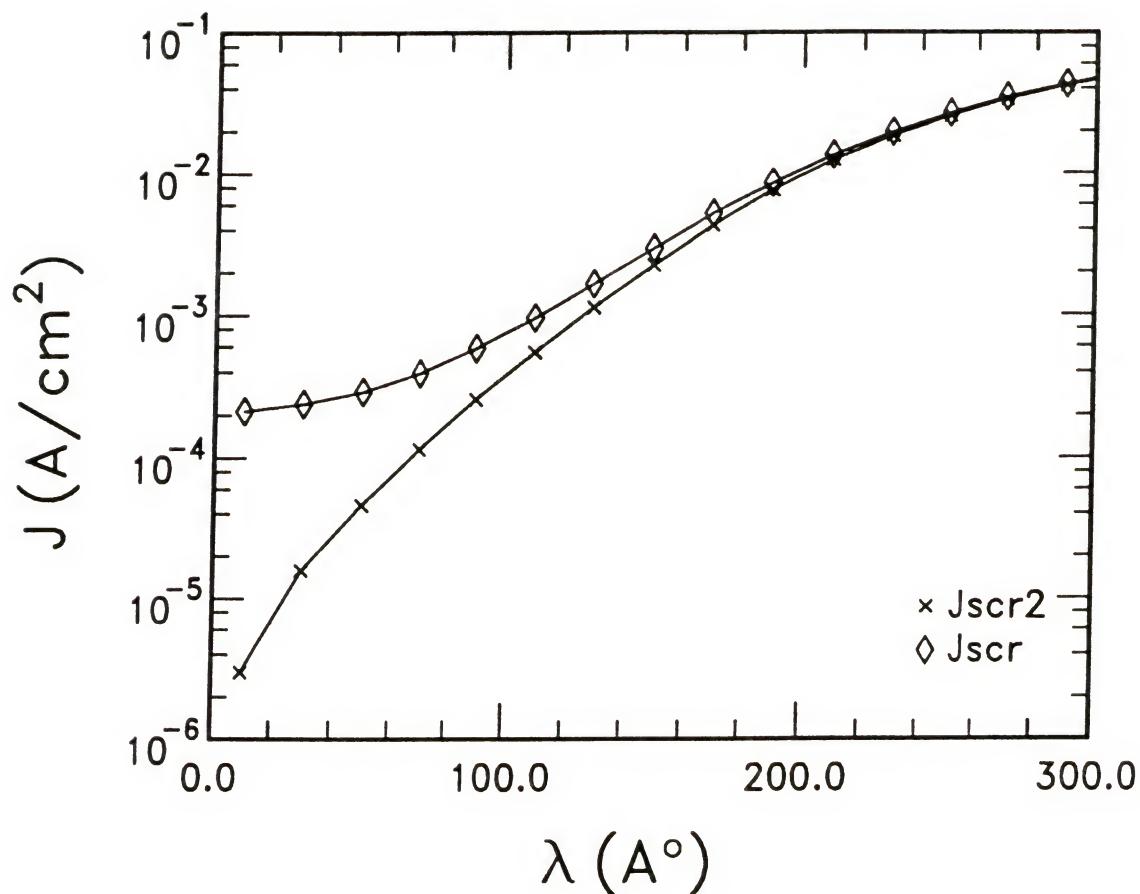
Fig. 6.4 Total SCR recombination current (J_{SCR}) and the recombination current in graded emitter layer (J_{SCR2}) as functions of the graded layer thickness for V_{BE} = 1.2 V.

constant at 1.2 V. We observe that above a grading width of about 150 A°, J_{SCR2} dominates the total recombination current density.

Fig. 6.5 shows the current gain β (=J_C/J_B) as a function of the collector current density. For all three grading widths, the current gain at higher current densities is seen to be approximately the same (for the parameters chosen here). This is because at these high current densities, the recombination in the quasi-neutral base region dominates the base current, and thus the SCR recombination characteristics become irrelevant. At lower current densities, the junction with a grading width of 100 A° gives the flattest characteristics. For the abrupt (1 A°) junction, the β is lower due to a lower collector current, which is caused by the conduction band peak. For the 300 A° grading width device, the β is lower (than the 100 A° device) due to higher recombination in the emitter SCR.

In Fig. 6.6 we compare the current gain obtained from the model with experimental results of Ma et al [Ma89]. The parameters are as in their paper, and those that are different from Figs. 6.2 and 6.3 are as follows: W_C = 0.3 μm, N_{DC} = 5 x 10^{16} /cm^3, N_{DE} = 5 x 10^{17} /cm^3, μ_n = 700 cm^2/V·s, and τ_n = τ_p = 30 ps. Excellent agreement is found between the experimental data and the analytical model.

In Fig. 6.7 we compare the base currents of an abrupt heterojunction transistor calculated from two different models: One (·) is the model proposed in this paper. The other (x) is calculated assuming ($E_{FN}-E_{FP}$) = qV_{BE}, which in effect means that we are ignoring the interdependence of the SCR recombination current and the current
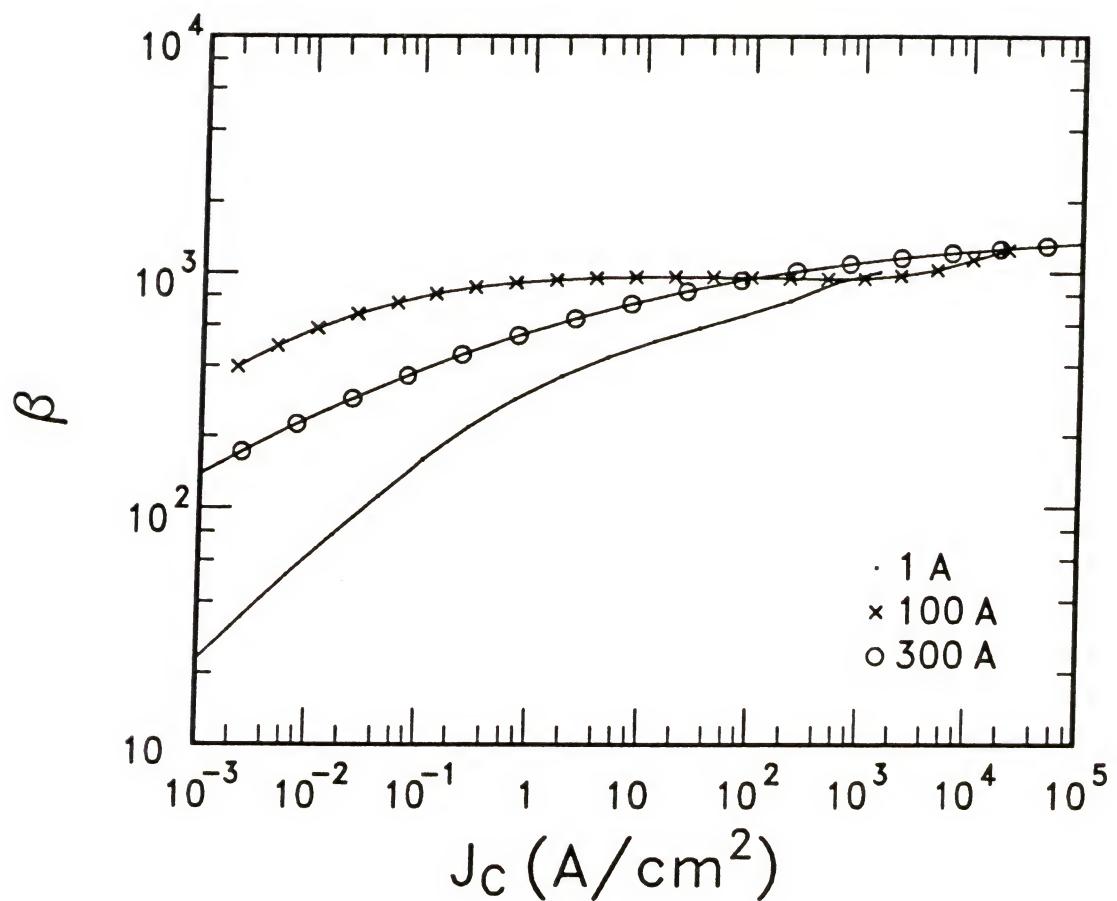
Fig. 6.5 Current gain vs. collector current characteristics for three graded layer thicknesses.
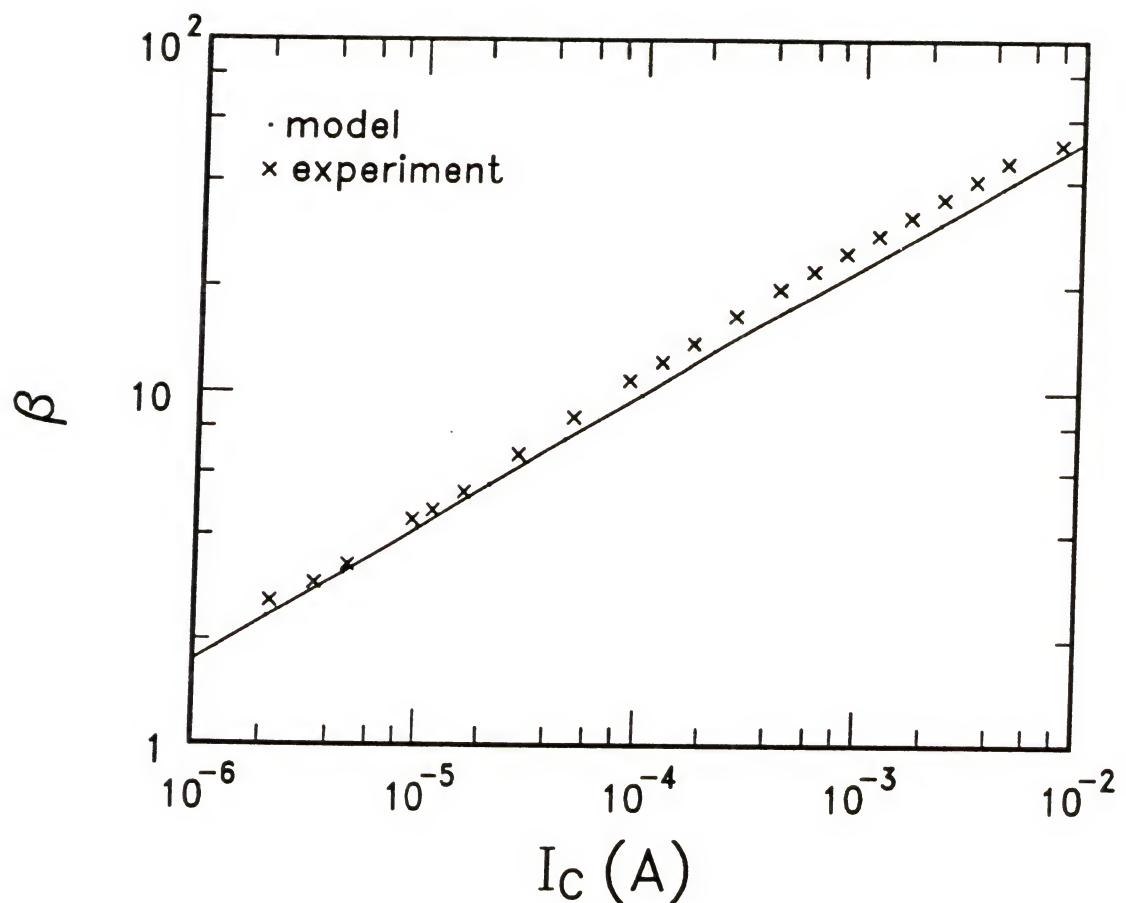
Fig. 6.6 Comparison of model with the experimental results of Ma et al
[Ma89].
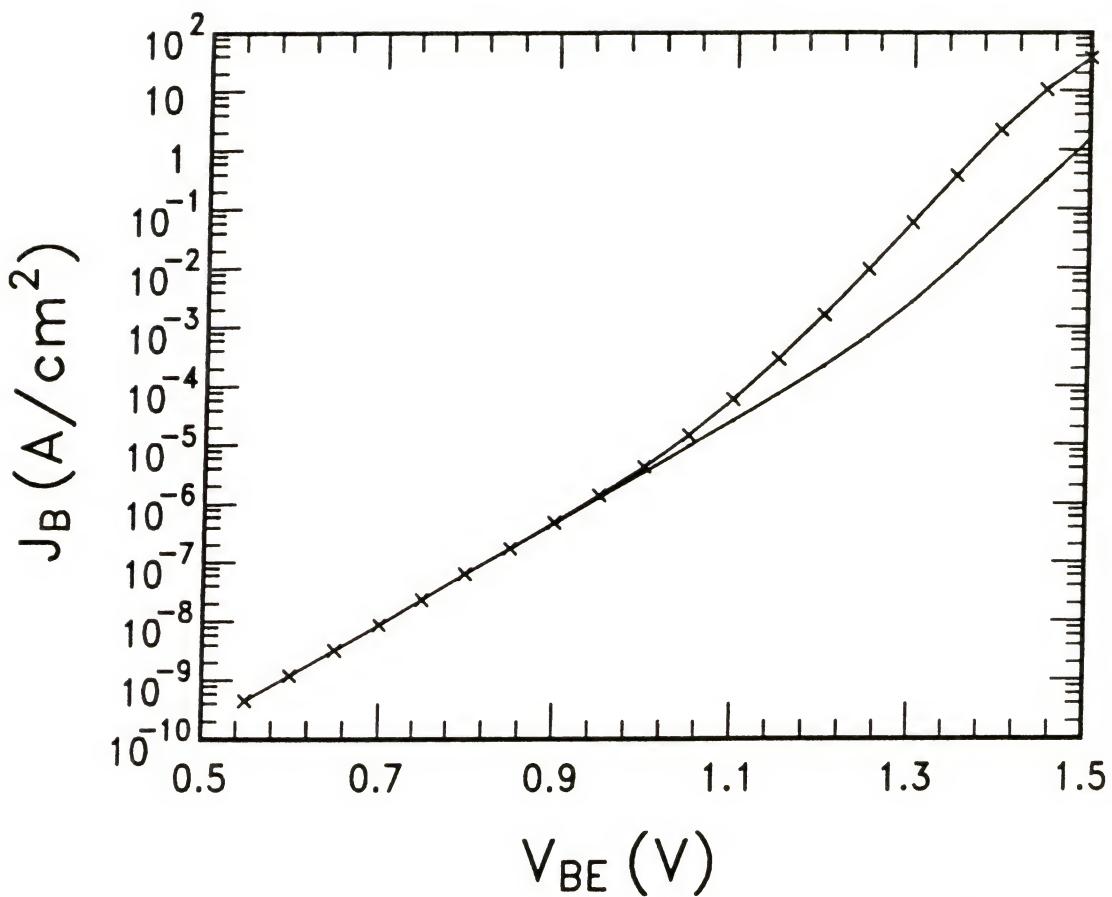
Fig. 6.7 Base current calculated using two different models.

through the base, as in a homojunction. We observe in the figure that the two currents differ by as much as an order of magnitude at higher biases. At low biases, the two currents are equal. This is because at these biases, most of the recombination occurs on the emitter side, which is not affected by the interdependence of currents.

6.5 Conclusions

In conclusion, we note that a unique interrelationship exists between the quasi neutral base transport current and the SCR recombination current, when a peak is present in one of the bands at the heterojunction interface. This interdependence is a result of the current flow over the peak being due to thermionic emission (or thermionic field emission). It is shown that ignoring this interdependence can cause the base current to be overestimated by as much as an order of magnitude.

Numerical results show that grading the emitter energy gap causes an increase in the recombination current in the graded region. This is because the intrinsic carrier density then becomes larger than what it would be in an abrupt heterojunction. This conclusion assumes that the recombination center density is independent of the presence or absence of energy gap grading.

An excellent agreement is found between experimental data [Ma89] and the model presented.

We also note the following points about heterojunctions.

In most heterojunctions, both surface and bulk recombination are found to play a significant role in determining the I-V characteristics.

Surface recombination, which is virtually non-existent in Si bipolar devices, plays a significant role in compound semiconductor HBT's [Hen77, Log77, Henr78a, Henr78b, Nak85, Hay87, Lyo89a, LeeW89] because good surface passivation materials which would reduce the surface state density have not yet been found, although a few ways of reducing surface recombination have been reported [Nak85, Tiwa87, San87, Nott88, Ma89]. Such recombination, occurring at surfaces that usually lie in the extrinsic base region [LeeW89], is dependent greatly on the fabrication technology and the device geometry. Not much attention has been given to modeling surface recombination [Sah62, Henr78a, Henr78b]. Getting rid of surface states is a technological problem, and we shall touch briefly upon this subject when we discuss heterojunction fabrication attempts in Sec. 9.1.

A source of additional recombination found in heterojunctions is at the metallurgical interface between the two dissimilar materials making up the heterojunction. This has been studied [Ahf73, Ett76, Jer76, Hen77, Henr78a, Henr78b, Kroe78, Mar83a, Mar83b, Mat83, Cap85, Jeon87], its effects have been modeled [Bar47, Kandi73, Hal75, Car79, Fah79, Kand79, Ovs80, Shi80, Shi81, Shi83], and its effective velocity measured [Zhu76, Ett77, Lan77, Nel78a, Nel78b, Abd79, McA82, End86]. But with the rapid improvements in fabrication technologies, interface recombination currents are becoming increasingly less important. For example the density of interface states in AlGaAs/GaAs heterojunctions fabricated by MBE or MOCVD, was found to be negligible [Lan77].

CHAPTER 7
MODELING QUASI-SATURATION PHENOMENA IN THE
COLLECTOR OF n/p/n GaAs TRANSISTORS

7.1 Introduction

High current densities in the collector of a bipolar transistor cause a significant voltage drop across the collector epitaxial layer (epi-layer) and give rise to the phenomenon of quasi-saturation (that is, the base-collector junction becomes forward biased in spite of a reverse bias applied at the terminals). These phenomena have been studied and modeled in detail for silicon bipolar transistors [Kir62, Bea68, Cla69, Pal69, Poo69, Whi69, Gra73, Bow73, Cho75, Gau75, Rey75, Kum75, Ald82, Kul85, Jeo87, Jeo89]. Of particular significance are the papers by Kirk [Kir62], who first modeled the phenomenon; Poon et al [Poo69], who did the first numerical simulation to demonstrate quasi-saturation; Bowler and Lindholm [Bow73], who gave a comprehensive qualitative description; Rey et al [Rey75], who derived analytical expressions for the current induced base thickness under all possible operating conditions; and Beale and Slatter [Bea68], Kull et al [Kul85] and Jeong and Fossum [Jeo87, Jeo89], all of whom derived the collector current (I_C) versus collector-emitter voltage (V_{CE}) characteristics in quasi-saturation. For GaAs heterojunction bipolar transistors (HBTs), on the other hand, little work has been done on understanding and modeling quasi-saturation phenomena [Lio89, Ryu90]. Nonetheless, numerical simulations have shown that base pushout does occur in GaAs HBTs, and causes current gain falloff at current densities above 10^5 A/cm^2 [Yos85]

121

if the collector is lightly doped (5×10^{16} /cm^3). Such situations are likely to prevail in power transistors [Kim85, She87, Hig88, Ued89, Unl89] which require a large collector breakdown voltage and thus a high collector resistivity. Thus in the pursuit of developing a comprehensive model for GaAs HBTs, it is essential that quasi-saturation phenomena occurring in the collector be modeled.

Quasi-saturation phenomena can be broadly divided into two regions: (i) ohmic, and (ii) nonohmic. This division refers to the nature of current flow in the collector epi-layer. When the current flow is ohmic, the drift velocity varies linearly with the electric field, and thus the characterization for both Si and GaAs transistors is identical. On the other hand, when the flow is nonohmic, the characterization becomes dependent on the exact nature of the drift velocity (v_d)- electric field (E) characteristics. Fig. 7.1 shows the v_d-E characteristics for electrons in silicon and gallium arsenide. It is the fact that the nonlinear portions of these characteristics are so different from each other that makes it necessary to derive different models for quasi-saturation phenomena in the two materials.

It is therefore the purpose of this chapter to derive a model that would take into account the unique nature of the v_d-E characteristics of electrons in GaAs in describing quasi-saturation phenomena.

The chapter is organized as follows. In the next section, a new model is derived for the current induced base thickness (W_{CIB}, defined below) for N/p/n GaAs transistors. In Sec. 7.3, I_C-V_{CE} characteristics
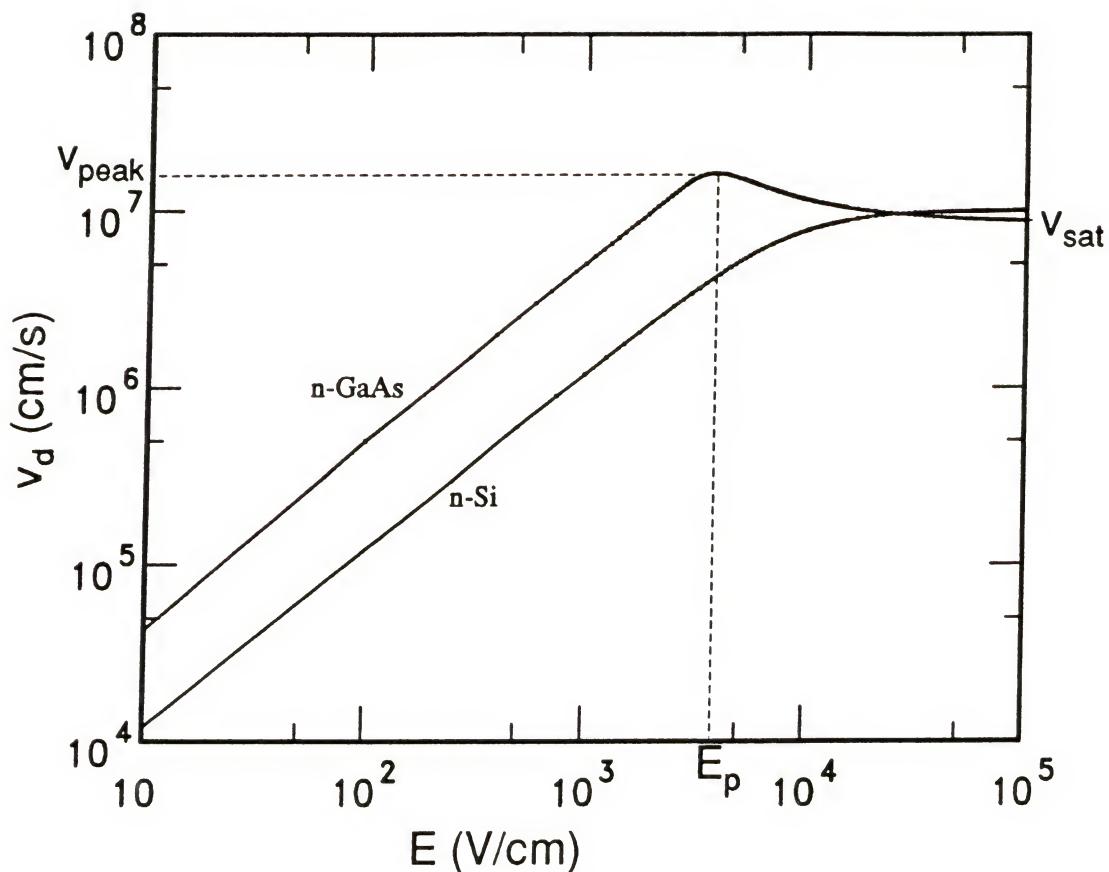
Fig. 7.1 Drift velocity vs. electric field characteristics of electrons in GaAs and Si.

are derived that take into account the high current phenomena within the collector. In Sec. 7.4, numerical results are presented.

We note some points here.

(i) We have modeled only a transistor with an n-type collector; and the model is not valid for a p-type collector because the v_d-E characteristics of holes in GaAs differ from those of electrons.

(ii) The model derived and discussed in Secs. 7.2 through 7.4 is valid for both homojunction and single heterojunction bipolar transistors, since it is only the collector region that is modeled here. But it is not valid for a double heterojunction transistor (DHBT); this is because if the base-collector junction is heterojunction, then minority carrier injection into the collector is greatly reduced, and base pushout is not likely to occur. We discuss a modeling approach for high current densities in DHBTs in Sec. 7.5. (iii) In an AlGaAs/GaAs heterojunction bipolar transistor, numerical simulations have indicated [Yos85] that the low mobility of electrons in AlGaAs causes high current effects to occur in the emitter before they occur in the collector, if the emitter is lightly doped. We assume in our analysis that such emitter effects are negligible.

7.2 Calculation of W_{CIB}

When quasi-saturation occurs, the base-collector junction is forward biased due to the voltage drop in the collector epi-layer caused by a high current flowing through it. Thus minority carriers are injected into the collector. If the level of injection of these minority carriers is high, then there exists a region adjacent to the

metallurgical base-collector junction, on the collector side, where both hole and electron densities are large compared to the collector doping density. In that region, recombination is significant, and current flow is by both diffusion and drift (region I in Fig. 7.2). The effect of such behavior on transistor parameters such as current gain (β) and cutoff frequency (f_T) can be modeled by treating the quasi-neutral base as if it had extended itself to include the region under high injection in the collector [Kir62, Whi69]. The region under high injection is therefore called the current induced base (CIB) region, the thickness of this region is called W_{CIB} (see Fig. 7.2), and this phenomenon is referred to as base pushout [Bow73]. Calculation of W_{CIB} is therefore an important aspect of modeling quasi-saturation phenomena.

When base pushout occurs, the collector layer can be divided into two regions (Fig. 7.2, cf. Fig. 1 of [Rey75]): region I, a quasi-neutral region in high injection, is adjacent to the base layer, and constitutes the "pushed-out" base; and region II, the region across which most of the potential drop occurs and has a substantial electric field across it. In GaAs collectors, current flow in region II is either ohmic or nonohmic depending on whether the electric field there is less or greater than E_p (see Fig. 7.1), respectively. We shall derive, in this section, expressions for the current induced base thickness - W_{CIB} - for both these cases.

We make the following assumptions in the analysis.
(1) The boundary region between regions I and II is narrow enough to be
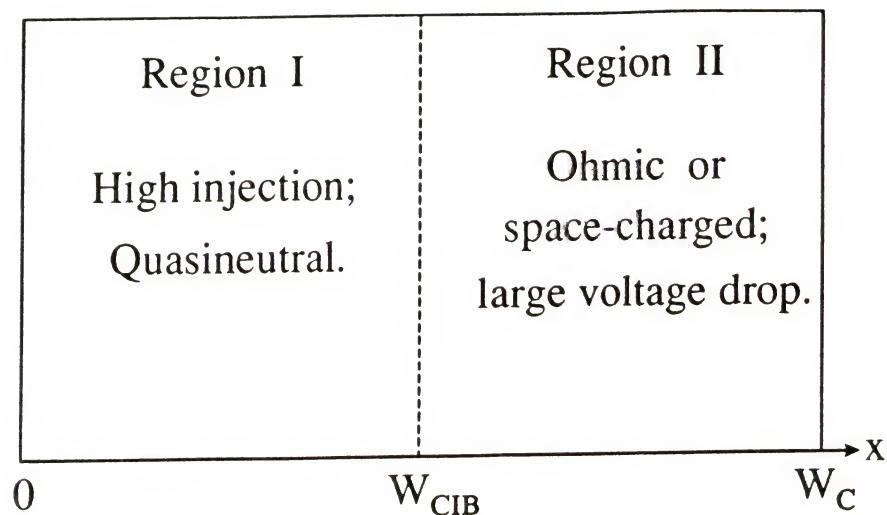
Fig. 7.2 Division of the collector layer into two regions under base pushout conditions.

considered abrupt. This has been borne out by numerical calculations [Poo69].

(2) In the boundary region between regions I and II, there will exist a space charge region; and the electric field will vary rapidly there [Poo69, Rey75]. Within this region, W_{CIB} is defined as the point at which $E = E_p$.

(3) The steady-state v_d - E characteristics of electrons in GaAs can be described by the empirical formulas [Chan86]

$$v_d = - \mu_{no} E \qquad \text{for } E \leq E_o$$

and

$$v_d = \frac{- \mu_{no} E}{\sqrt{1 + \left(\dfrac{E_o - E}{E_c} \right)^2}} \qquad \text{for } E \geq E_o \qquad (7.2\text{-}1)$$

where

$$E_o = 1/2 \ [\ E_p - (\ E_p^2 - 4 \ E_c^2 \ )^{1/2} \ ] \qquad (7.2\text{-}1a)$$

E_p = field at which the velocity peaks

$$E_c = v_{sat} \ / \ \mu_{no} \qquad (7.2\text{-}1b)$$

v_{sat} = saturation velocity at high fields (Fig. 7.1)

μ_{no} = low field mobility

E is assumed to be negative (as it will be in the collector). There has been another empirical formula suggested for the v_d-E characteristics by Xu and Shur [Xu81]; the only reason for choosing the formula of Eqn. (7.2-1) is that it makes analytical integration of Poisson's equation possible.

The appropriateness of the use of steady-state v_d-E characteristics for modeling the collector layer has been questioned by Liou et al [Lio90]. Their argument was that the voltage drop across the space charge region, and its thickness were too small to allow electrons to get transferred to the higher L and X valleys; and thus the steady-state characteristics, which assume electron transfer to higher valleys, were inappropriate. But, as we discussed in Sec. 4.7 (and will discuss further in Sec. 9.3), Monte Carlo simulations [Roc88, Malo77, Gru80] investigating velocity overshoot in the base-collector space charge region (SCR) show that valley transfer does occur in all the transistors considered in such simulations. We therefore assume that the steady-state characteristics are appropriate. We thus also assume that the effect of velocity overshoot is negligible.

(4) Region II is in low injection. This follows from the way regions I and II are defined, since region I is the part of the collector that is in high injection. This allows us to neglect the hole density in region II while solving Poisson's equation there.

(5) The diffusion current is negligible in region II, since the minority hole density in this region is small.

(6) When the transistor is operating in base-pushout, the potential drop across the base-collector junction space charge layer is negligible compared to the drop across the collector epi-layer. This is because for base-pushout to occur, this junction must be heavily forward biased, and thus almost all of the potential drop will occur across the collector epi-layer.

7.2.1 Ohmic Conduction

Here, since the electric field in region II is less than E_p, the mobility is assumed constant; the collector current then is written as,

$$J_C = q \ \mu_n \ N_{DC} \ \frac{V_{CB}+V_{bi}}{W_C-W_{CIB}} \tag{7.2-2}$$

where V_{CB} is the applied voltage, V_{bi} is the built-in voltage, and we have neglected the diffusion current. Rearranging yields [Rey75]

$$W_{CIB} = W_C \left[1 - \frac{q\mu_n N_{DC}}{J_C} \frac{V_{CB}+V_{bi}}{W_C} \right] \tag{7.2-3}$$

The result is identical to that for silicon.

7.2.2 Nonohmic Conduction

The maximum current density that a neutral collector region can support is

$$J_{Cmax} = q \ N_{DC} \ v_{peak}$$

where v_{peak} is the peak electron velocity (Fig. 7.1). For current densities greater than this, electrons must be injected into the collector [Bow73]. This would cause space charge to accumulate in region II; thus the electric field, which was E_p at J_{Cmax}, must now increase to satisfy Poisson's equation (Eqn. (7.2-4) below). Thus, to analyze this case, we must solve Poisson's equation.

We write Poisson's equation in region II as

$$\frac{dE}{dx} = \frac{q}{\varepsilon} \ (\ N_{DC} - n \ ) \tag{7.2-4}$$

where we have used assumption (4). With assumption (5) the current density will be

$$J_C = - q n v_d \tag{7.2-5}$$

Combining (7.2-4) and (7.2-5) gives

$$\frac{dE}{dx} = \frac{q}{\varepsilon} \left(N_{DC} + \frac{J_C}{qv_d} \right)$$

$$= \frac{qN_{DC}}{\varepsilon} \left(\frac{v_d + J_C/qN_{DC}}{v_d} \right)$$

Integration gives

$$\frac{qN_{DC}}{\varepsilon} \int dx + C_E = \int \frac{v_d}{v_d + J_C/qN_{DC}} \, dE$$

where C_E is the constant of integration. Let

$$v_1 \overset{\Delta}{=} \frac{J_C}{qN_{DC}}$$

Then

$$\frac{qN_{DC}}{\varepsilon} \int dx + C_E = \int \frac{v_d}{v_d + v_1} \, dE \overset{\Delta}{=} A_E \tag{7.2-6}$$

The task now is to evaluate the integral for A_E. Remembering Eqn. (7.2-1) for v_d, we use the substitution,

$$(E_0-E)/E_c = \sinh w \tag{7.2-7}$$

to yield

$$A_E = - E_c \int \frac{- \mu_{no} (E_0 - E_c \sinh w) / \cosh w}{-\mu_{no}E_0 + \mu_{no}E_c \sinh w + v_1 \cosh w} \, dw$$

Now letting

$$\tanh (w/2) = z \tag{7.2-8}$$

and

$$\frac{v_d}{\mu_{no}} = \frac{J_C}{qN_D\mu_{no}} \overset{\Delta}{=} E_d \tag{7.2-9}$$

gives

$$A_E = -2E_c \int \frac{(E_o z^2 + 2E_c z - E_o) (1 + z^2)}{(1+z)^2 (1-z)^2 [(E_o+E_d)z^2 + 2E_c z + E_d-E_o]} dz \qquad (7.2\text{-}10)$$

The integrand can be expanded into partial fractions as

$$\text{Integrand} = \frac{A}{(1+z)^2} + \frac{B}{(1+z)} + \frac{C}{(1-z)^2} + \frac{D}{(1-z)} + \frac{Fz + G}{[(E_o+E_d)z^2+2E_c z+E_d-E_o]}$$

where

$$A = \frac{E_c}{2(E_c-E_d)} \qquad B = \frac{-E_o E_d}{2(E_d-E_c)^2} \qquad C = \frac{E_c}{2(E_c+E_d)} \qquad D = \frac{-E_o E_d}{2(E_d+E_c)^2}$$

$$F = (E_o+E_d) (D-B) \qquad G = -E_o + (A+B+C+D) (E_o-E_d)$$

Integrating these partial fractions gives the final integral as

$$\frac{A_E}{-2E_c} = \frac{-A}{z+1} + B \ln|z+1| - \frac{C}{z-1} - D \ln|z-1|$$
$$+ \frac{F}{E_o+E_d}[\frac{1}{2}\ln|z^2+2bz+c|+(h-b)\cdot T6] \qquad (7.2\text{-}11)$$

where

$$b = E_c / (E_o+E_d) \qquad c = (E_d-E_o) / (E_d+E_o)$$
$$h = G / F \qquad\qquad\qquad\qquad\qquad\qquad\qquad (7.2\text{-}11a)$$

and

$$T6 = \frac{1}{\sqrt{c-b^2}} \tan^{-1}\left[\frac{z+b}{\sqrt{c-b^2}}\right] \qquad \text{if } c-b^2 > 0$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (7.2\text{-}11b)$$

$$= \frac{1}{2\sqrt{b^2-c}} \ln \left|\frac{z+b-\sqrt{b^2-c}}{z+b+\sqrt{b^2-c}}\right| \qquad \text{if } c-b^2 < 0$$

From Eqns. (7.2-7) and (7.2-8) we get for z

$$z = \frac{\sqrt{1 + [(E_o-E)/E_c]^2} - 1}{(E_o-E)/E_c} \qquad (7.2-12)$$

Now if the electric field at $x=W_C$ is defined to be E_w, then from Eqn. (7.2-6) the following relation is obtained:

$$W_{CIB} = (\varepsilon/qN_{DC}) [A_E(E_p) - A_E(E_w)] + W_C \qquad (7.2-13)$$

assuming that $E(x=W_{CIB}) = E_p$. Thus once E_w is known, W_{CIB} can be obtained from the above equation. To obtain E_w, we must integrate Poisson's equation once more to relate the electric field to the applied voltage. This we do now.

By definition

$$V = - \int E \, dx \equiv - \int \frac{E}{dE/dx} \, dE = - \int \frac{E}{(qN_{DC}/\varepsilon) \, (v_d+v_1)/v_d} \, dE \quad (7.2-14)$$

where we have used Eqn.s (7.2-4,5,6). Hence

$$- \frac{qN_{DC}}{\varepsilon} V = \int \frac{Ev_d}{v_d + v_1} \, dE$$

Using the same substitutions (7.2-7) and (7.2-8), we get the integral

$$- \frac{qN_{DC}}{\varepsilon} V = 2E_c \int \frac{(E_o z^2 + 2E_c z - E_o)^2 (1 + z^2)}{(1+z)^3 (1-z)^3 [(E_o+E_d)z^2 + 2E_c z + E_d-E_o]} \, dz$$

Letting

$$\eta_c \overset{\Delta}{=} \frac{E_c}{E_o} , \qquad \eta_d \overset{\Delta}{=} \frac{E_d}{E_o}$$

the above integrand can be expanded into partial fractions as

$$\text{Integrand} = \frac{A_1}{(1+z)^3} + \frac{A_2}{(1+z)^2} + \frac{A_3}{(1+z)} + \frac{A_4}{(1-z)^3} + \frac{A_5}{(1-z)^2} + \frac{A_6}{(1-z)}$$

$$+ \frac{A_7 z + A_8}{[\ (1+\eta_d)z^2 + 2\eta_c z + \eta_d - 1\ ]}$$

where

$$A_1 = \frac{\eta_c^2}{2(\eta_d - \eta_c)}\ , \qquad A_2 = \frac{\eta_c(4\eta_d - 2\eta_c - \eta_c\eta_d + \eta_c^2)}{4(\eta_d - \eta_c)^2}$$

$$A_3 = \frac{-\eta_c^2\eta_d^2 + \eta_c^3\eta_d + 2\eta_d^2}{4(\eta_d - \eta_c)^3}\ ,$$

$$A_4 = \frac{\eta_c^2}{2(\eta_d + \eta_c)}\ , \qquad A_5 = \frac{-\eta_c(4\eta_d + 2\eta_c + \eta_c\eta_d + \eta_c^2)}{4(\eta_d + \eta_c)^2}$$

$$A_6 = \frac{-\eta_c^2\eta_d^2 - \eta_c^3\eta_d + 2\eta_d^2}{4(\eta_d + \eta_c)^3}\ ,$$

$$A_7 = (\ 1 + \eta_d\ )\ (\ A_6 - A_3\ )\ , \qquad A_8 = 1 + (1 - \eta_d)\ (A_1 + A_2 + A_3 + A_4 + A_5 + A_6)$$

Evaluating the integral gives the final result as

$$- \frac{qN_{DC}}{\varepsilon} V(E) = \frac{-A_1}{2(1+z)^2} - \frac{A_2}{(1+z)} + A_3\ \ln|1+z| + \frac{A_4}{2(1-z)^2} - \frac{A_5}{(z-1)} -$$

$$- A_6\ \ln|z-1| + \frac{A_7}{1+\eta_d}\left[\ \frac{1}{2}\ \ln|z^2 + 2p_1 z + p_2| + T7\ \right]\ (7.2\text{-}15)$$

where

$$p_1 = \frac{\eta_c}{1 + \eta_d}\ , \quad p_2 = \frac{\eta_d - 1}{\eta_d + 1}\ , \quad z_1 = A_8\ /\ A_7$$

and

$$T7 = \frac{z_1 - p_1}{\sqrt{p_2 - p_1^2}} \tan^{-1}\left[\frac{z + p_1}{\sqrt{p_2 - p_1^2}}\right] \qquad \text{if } p_2 - p_1^2 > 0$$

$$= \frac{z_1 - p_1}{2\sqrt{p_1^2 - p_2}} \ln\left|\frac{z + p_1 - \sqrt{p_1^2 - p_2}}{z + p_1 + \sqrt{p_1^2 - p_2}}\right| \qquad \text{if } p_2 - p_1^2 < 0$$

From assumption (6) the boundary condition across the collector layer can be written as

$$V(E_w) - V(E_p) = V_{CB} + V_{bi} \qquad (7.2\text{-}16)$$

where V_{CB} is the applied bias, V_{bi} is the built-in voltage, and $V(E_w)$ and $V(E_p)$ are respectively the voltages at $x = W_C$ and $x = W_{CIB}$. This equation can be solved (numerically) to give E_w, which then can be used to find W_{CIB} using Eqn. (7.2-13). We thus have obtained W_{CIB} as a function of the collector current density and the base collector voltage.

7.3 I_C-V_{CE} Characteristics in Quasi-saturation

Quasi-saturation affects the I_C-V_{CE} characteristics of the transistor by causing the base-collector (B-C) junction to remain forward biased even when a reverse bias exists at the terminals. This has been modeled in Si transistors by Beale and Slatter [Bea68], Kull et al [Kul85] and Jeong and Fossum [Jeo89]. In this section we will derive equations to describe the I_C-V_{CE} characteristics of a GaAs transistor in quasi-saturation. The equations obtained will not be explicit relationships, but rather a set of transcendental equations; they would augment the model of chapter 5 quite naturally. The combined set of

equations would need to be solved numerically. We will consider the cases of low and high injection, and ohmic and nonohmic current flow in the collector.

7.3.1 High Injection in the Collector

In this regime, the collector can be divided into two regions, as we discussed earlier (see Fig. 7.2). The approach we use is to determine the currents flowing through regions I and II, equate these two to obtain the current flowing through the collector, and then equate this latter current with the expression for the current through the base to obtain a transcendental equation, which when solved would yield the I-V characteristics.

7.3.1.1 Ohmic conduction

To obtain an analytical model, we make the following assumptions for region I (as were made in the models for silicon transistors [Bea68, Kul85]):

(1) $p \simeq n$, since the region is assumed to be in high injection.

(2) Negligible minority (hole) current compared to the majority (electron) current. The justification for this assumption is as follows. The hole drift and diffusion currents are in opposite directions, while the electron drift and diffusion currents are in the same direction. Thus, since $p \simeq n$, the net hole current is small compared to the net electron current.

(3) The recombination current is negligible compared to the electron current.

The current flowing in region I then is

$$J_C \simeq J_N = q \ \mu_n \ n \ E + q \ D_n \ dn/dx \tag{7.3-1}$$

Also

$$J_P = q \ \mu_p \ p \ E - q \ D_p \ dp/dx$$

Solving this for E and substituting in Eqn. (7.3-1) yields,

$$J_C = q \ D_n \left(\frac{n}{p} \frac{dp}{dx} + \frac{dn}{dx} \right) + \frac{\mu_n}{\mu_p} \ J_P$$

Assumptions (1) and (2) then give

$$\therefore \ J_C \simeq 2 \ q \ D_n \ \frac{dp}{dx} \tag{7.3-2}$$

Integrating Eqn. (7.3-2) from x=0 to W_{CIB} gives

$$\int_0^{W_{CIB}} J_C \ dx = 2 \ q \ D_n \int_0^{x=W_{CIB}} dp$$

$$J_C = \frac{2 \ q \ D_n}{W_{CIB}} \ [\ p(W_{CIB}) - p(0) \ ] \tag{7.3-3}$$

Let $\quad J_C \overset{\Delta}{=} \dfrac{K_1}{W_{CIB}}$ $\qquad\qquad\qquad\qquad$ (7.3-3a)

Now the same J_C must flow through region II. From Eqn. (7.2-2)

$$J_C = - \ q \ \mu_n \ N_{DC} \ \frac{V_{CC'}}{W_C - W_{CIB}} \overset{\Delta}{=} \frac{K_2}{W_C - W_{CIB}} \tag{7.3-3b}$$

Note that we write the voltage drop across region II as $V_{CC'}$ rather than $(V_{CB} + V_{bi})$; we shall discuss this shortly.

Equating the two expressions for J_C and rearranging gives

$$\frac{K_1}{W_{CIB}} = \frac{K_1 + K_2}{W_C}$$

Hence $\quad J_C = \dfrac{K_1 + K_2}{W_C}$

$$= \frac{q\ D_n}{W_C} \left\{ 2\ [p(W_{CIB})-p(0)] - \frac{N_{DC}}{V_T}\ V_{CC'} \right\} \qquad (7.3\text{-}4)$$

where $V_T = kT/q$. Since region I is in high injection,

$$p(0) = n_i^2/N_{DC}\ \exp(qV_{BC'}/2kT) \qquad (7.3\text{-}5)$$

where $V_{BC'}$ is the forward bias across the base-collector junction. Also, defining a point C' as the edge of the SCR on the collector side, the voltage drops in the region would be related as

$$V_{CB} + V_{bi} = V_{CC'} + V_J \qquad (7.3\text{-}6a)$$

and $\qquad V_J = V_{bi} - V_{BC'} \qquad (7.3\text{-}6b)$

where V_J is the electrostatic potential drop across the junction space charge region. Assuming $p(W_{CIB}) = p_0 = n_i^2/N_{DC}$, Eqn. (7.3-4) can be written as

$$J_C = -\frac{q\ D_n}{W_C} \left\{ \frac{2n_i^2}{N_{DC}} \left[\exp\left(\frac{V_{BC'}}{2V_T} \right) - 1 \right] + N_{DC}\ \frac{V_{CC'}}{V_T} \right\} \qquad (7.3\text{-}7)$$

The above characterizes the current through the collector in ohmic quasi-saturation. This now must be combined with the current characterization in the base to yield the terminal I-V characteristics.

For simplicity we assume a GaAs homojunction bipolar transistor with a uniform, highly doped base. Then we can use the simple Ebers-Moll model for current flow through the base [Sze81]. Thus we write the electron current injected into the collector as

$$J_C = q\ \frac{D_n\ n_i^2}{N_A\ L_n\ \sinh(W_B/L_B)} \left\{ \exp\left[\frac{qV_{BE}}{kT} \right] - 1 - \coth\left(\frac{W_B}{L_B} \right) \left(\exp\left[\frac{qV_{BC'}}{kT} \right] - 1 \right) \right\}$$

$$(7.3\text{-}8)$$

Equating (7.3-7) and (7.3-8) will give a transcendental equation for $V_{BC'}$. Solving for $V_{BC'}$ and substituting in (7.3-7) or (7.3-8) will yield J_C as a function of V_{BE} and V_{BC}. We emphasize that Eqn. (7.3-8) is the simplest possible model for base transport. More complex models, such as the Gummel-Poon model [Gum70b], or models for the heterojunction bipolar transistor that we reviewed in Secs. 4.2 through 4.5, or the model proposed in chapter 5, can equally well replace Eqn. (7.3-8). The sole reason for using Eqn. (7.3-8) is to obtain computational simplicity, since the more complex the model for J_C, the more involved the numerical calculations become.

7.3.1.2 Nonohmic conduction

The equations here are more complex than in the ohmic case. To begin with, we combine equations (7.2-13), (7.3-3) and (7.3-5) to obtain

$$J_C = \frac{2\ q\ D_n\ n_i^2/N_{DC}\ [\ \exp(qV_{BC'}/2kT)\ -\ 1\ ]}{\{\ (\varepsilon/qN_{DC})\ [\ A_E(E_p)\ -\ A_E(E_w)\ ]\ +\ W_C\ \}} \qquad (7.3-9)$$

Equating (7.3-8) with (7.3-9) yields a transcendental equation in two variables: $V_{BC'}$ and E_w. The other equation needed to solve the problem is Eqn. (7.2-16):

$$V(E_w)\ -\ V(E_p)\ =\ V_{CB}\ +\ V_{bi} \qquad (7.2-16)$$

Thus we have two equations in two unknowns, and they must be solved simultaneously to yield the final I-V characteristics. We note that in the last two equations, A_E and V are both functions of $J_N(=J_C)$; we use Eqn. (7.3-8) for J_C in these functions.

7.3.2 Low Injection in the Collector

If quasi-saturation occurs while the collector is in low injection, then there is no base pushout, so that W_{CIB} is zero (or, region I of Fig. 7.2 is nonexistent). For deriving the terminal I-V characteristics, we must once again consider two cases.

7.3.2.1 Ohmic conduction

If the current flow is ohmic in the collector, then we use Eqn. (7.3-3b) with $W_{CIB} = 0$ and Eqns. (7.3-6) to write

$$J_C = - q \mu_n N_{DC} \frac{V_{CC'}}{W_C} \qquad (7.3-10)$$

Then equate this with Eqn. (7.3-8) to obtain a transcendental equation for $V_{BC'}$. The solution of this equation, when substituted back into Eqn. (7.3-8) will yield the I-V characteristics.

7.3.2.2 Nonohmic conduction

Here we solve Eqn.s (7.2-13) and (7.2-16), with $W_{CIB} = 0$, for E_w and $V_{BC'}$, with Eqn. (7.3-8) for J_C. The solution of this system of equations will yield the desired characteristics.

7.4 Numerical Results

The experimental and numerical data reported in the literature on quasi-saturation phenomena in GaAs HBTs are never described in sufficient detail to facilitate comparison with the model developed here. Thus we will just examine some numerical results obtained from our model on their own.

In all the figures, the doping densities in the base and collector are respectively 10^{18} /cm^3 and 10^{15} /cm^3. The low field

electron mobility in the collector is assumed to be 4250 V/cm^2, and the collector layer thickness to be 3 μm (unless specified otherwise).

Figures 7.3(a) and 7.3(b) show the electric field profiles in the collector under nonohmic base pushout conditions. Fig. 7.3(a) shows the profiles for four different collector current densities, with the base-collector voltage kept constant. Since V_{BC} is held constant, the area under each of the curves must remain constant. Thus as the current density is increased, the profile is "pushed" away from the metallurgical junction, which is at x = 0. This behavior is similar to that in a silicon transistor. The difference here is the significant nonlinearity of the profile, which arises due to the peculiar v_d-E characteristics of electrons in GaAs. In silicon, once nonohmic current flow sets in, the drift velocity saturates, so that the current density in the collector is simply written as $J_C = q\ v_{sat}\ (N_D+\Delta n)$, where Δn is the excess electron density. Thus for a particular collector current density, Δn is a constant. Also, from Poisson's equation, $dE/dx \simeq q\Delta n/\varepsilon$. Hence dE/dx is a constant, and the electric field profile in the collector is linear (see e.g. [Rey75]). In GaAs, on the other hand, the drift velocity reaches a peak at $E=E_p$. As the current is increased beyond the peak current (J_{Cpeak}), the current density can be written as $J_C = q\ v_d(E)\ (N_D+\Delta n)$. Thus Δn in this case does not remain constant but must increase with E, since v_d decreases with E. Hence dE/dx also must increase with E (and hence x) rather than remaining constant. Hence the nonlinearity in the electric field profile. We observe though that at
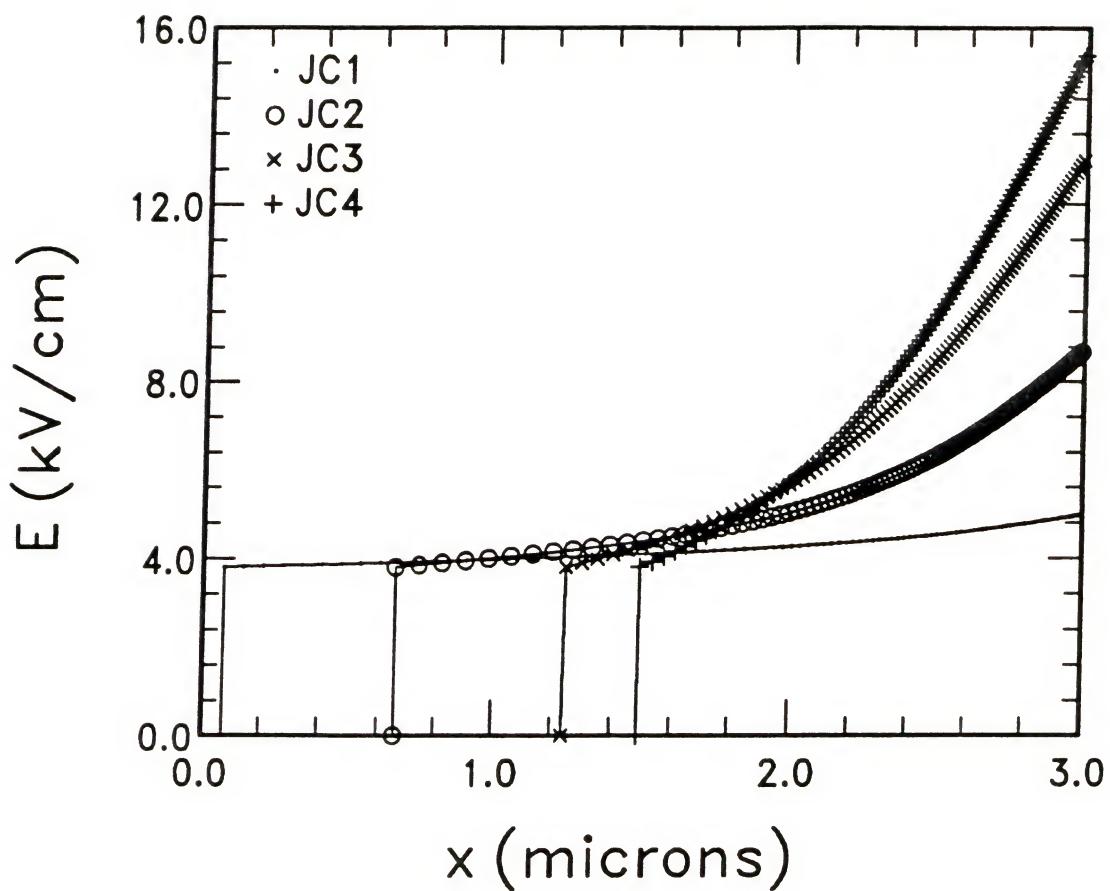
Fig. 7.3 (a) Electric field profile in the collector under nonohmic quasi-saturation: JC1 = 2620 A/cm^2, JC2 = 2690 A/cm^2, JC3 = 2890 A/cm^2, and JC4 = 3090 A/cm^2.
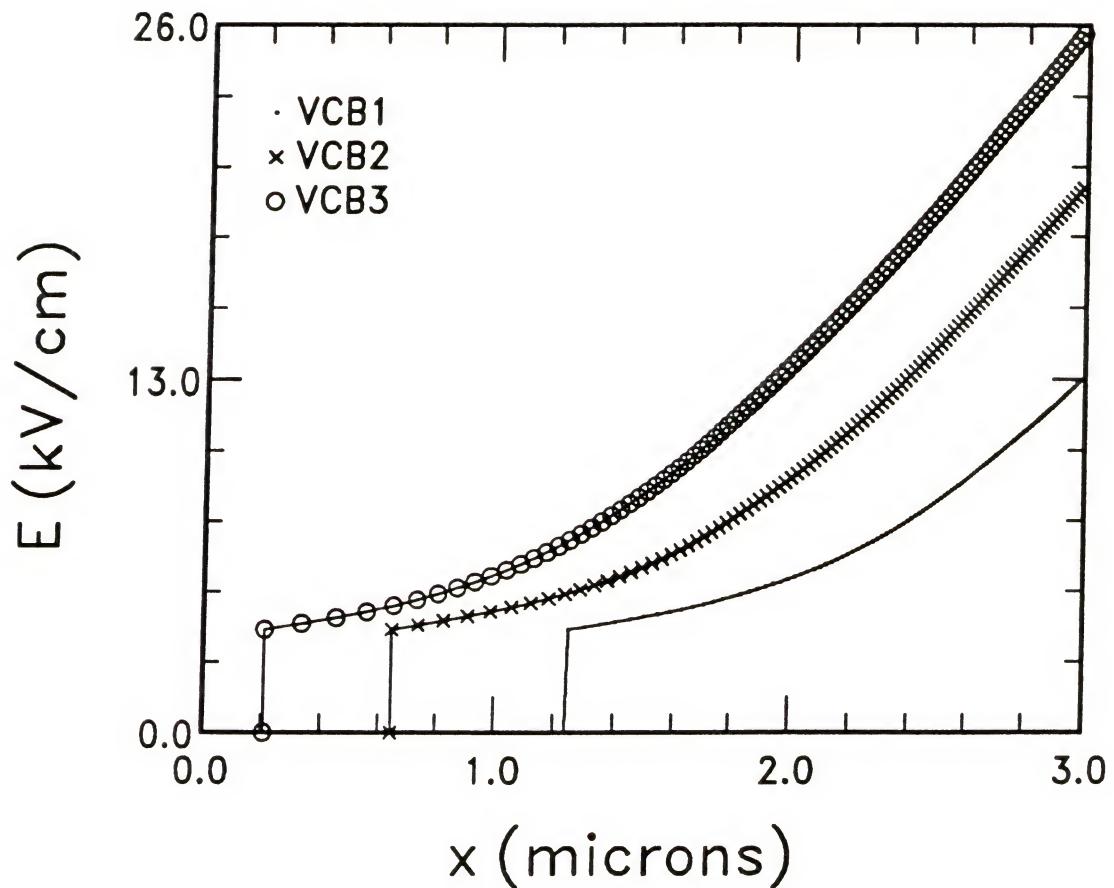
Fig. 7.3 (b) Electric field profile in the collector under nonohmic
          quasi-saturation: VCB1 = 0 V, VCB2 = 10 V, VCB3 = 20 V.

very high fields, the profile does start becoming linear since the drift velocity then starts saturating.

Fig. 7.3(b) shows the electric field profile for three different collector-base (reverse) biases. With increasing reverse bias, the area under the profile must increase, as is seen in the figure.

Fig. 7.4(a) shows the J_C-V_{CE} characteristics for the same transistor for three different base current densities. The peak ohmic collector current density (J_{cpeak}) for this transistor is about 2600 A/cm^2. Below this value, the characteristics exhibit ohmic quasi-saturation. Above this value, the characteristics show the effects of nonohmic quasi-saturation. Fig. 7.4(b) shows the J_C-V_{CE} characteristics for $J_B = 15$ A/cm^2 along with a profile that neglects all quasi-saturation effects. The difference, as indicated by the figure, is substantial. We note however, that it is for the purpose of illustration that such a drastic difference is shown here. In actual transistors, the collector thicknesses are much smaller than 3 μm, and hence so is the effect of quasi-saturation (the smaller the collector layer thickness, the smaller will be the voltage drop across the collector layer, and hence smaller will be the effect of quasi-saturation). At the same time, the figure indicates that in power transistors, where considerations of a large breakdown voltage necessitate a large, lightly doped collector layer, quasi-saturation will play an important role in the device characteristics.
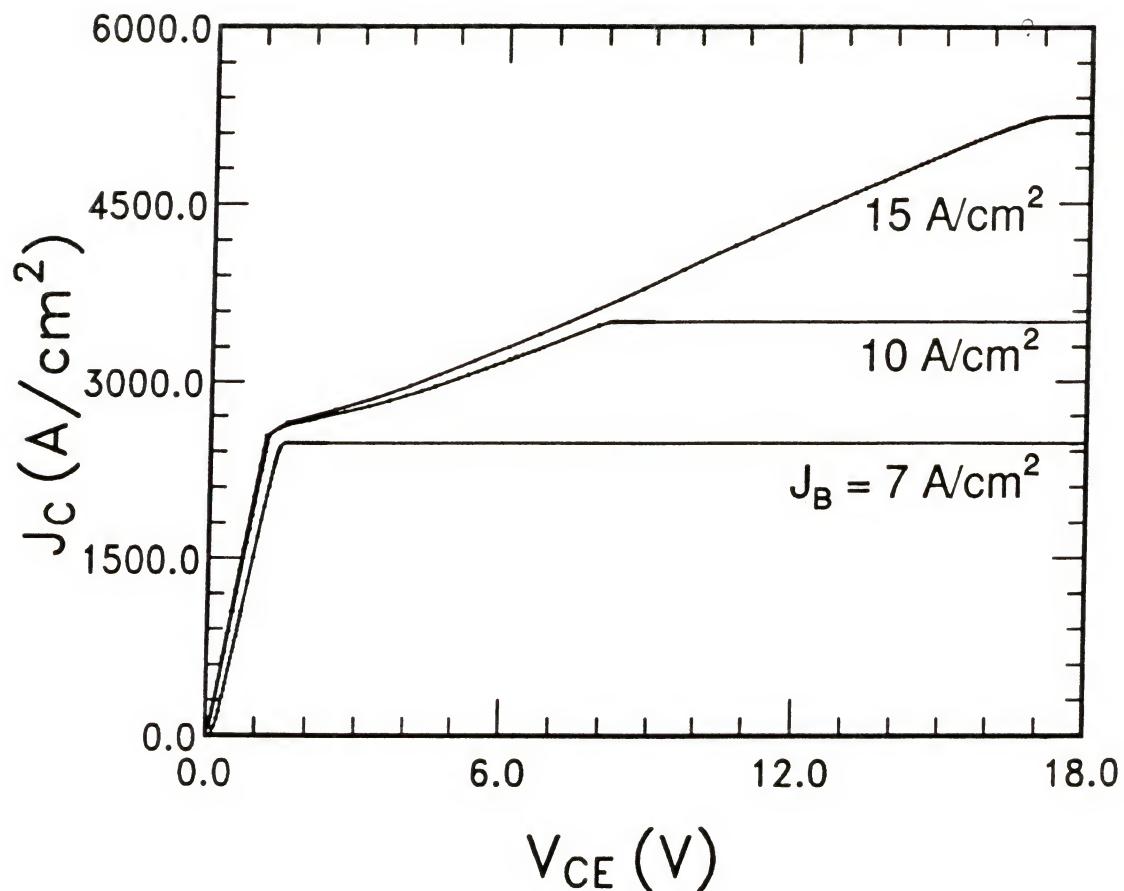
Fig. 7.4(a) J_C - V_{CE} characteristics showing the effects of quasi-saturation for three different base current densities.
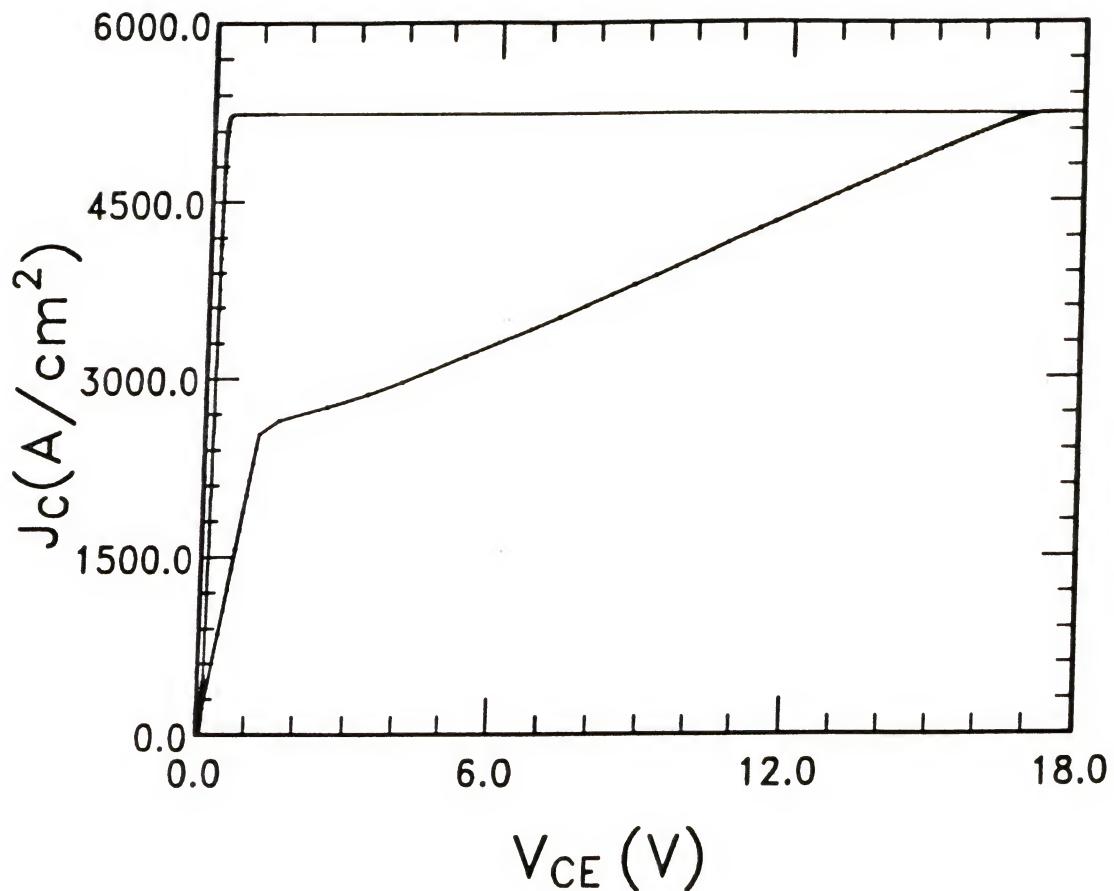
Fig. 7.4(b) J_C - V_{CE} characteristics showing the effects of quasi-saturation for the same base current density as in Fig. 7.4(a) but with and without including the effects of quasi-saturation in the calculations.

Finally, Fig. 7.5 shows the current gain (β) as a function of the collector current density, for three different collector layer thicknesses. A simple formula for the current gain is used:

$$1/\beta = 1/\beta_1 + 1/\beta_2$$

where

$$\beta_1 = J_C / J_{SCR}$$

$$\beta_2 = 4 \, L_n^2 / (\, 2W_B^2 + W_{CIB}^2 \,)$$

where J_{SCR}, the space charge region recombination current, is calculated from the model of chapter 6. We use L_n = 3.3 μm and W_B = 0.1 μm; all lifetimes are taken to be 1 ns. An $Al_yGa_{1-y}As$/GaAs HBT is assumed with a graded layer thickness of 500 A°, and y = 0.2. We observe that as the collector layer thickness is increased, β falls off earlier, and the fall-off is sharper. For all three curves, β tends to saturate after an initial sharp fall-off. This is because once W_{CIB} reaches W_C, it cannot increase further, and thus β also does not decrease further.

7.5 High Current Densities in Double Heterojunction Bipolar Transistors

In the previous sections of this chapter, we studied, analytically and numerically, the effects of high current densities flowing through the collector of a GaAs bipolar transistor in which the base-collector junction was a homojunction. In this section we look at the effects of high current densities in a double heterojunction bipolar transistor (DHBT).

The presence of a base-collector heterojunction, in an N/p/N transistor, would drastically reduce the injection of holes into the
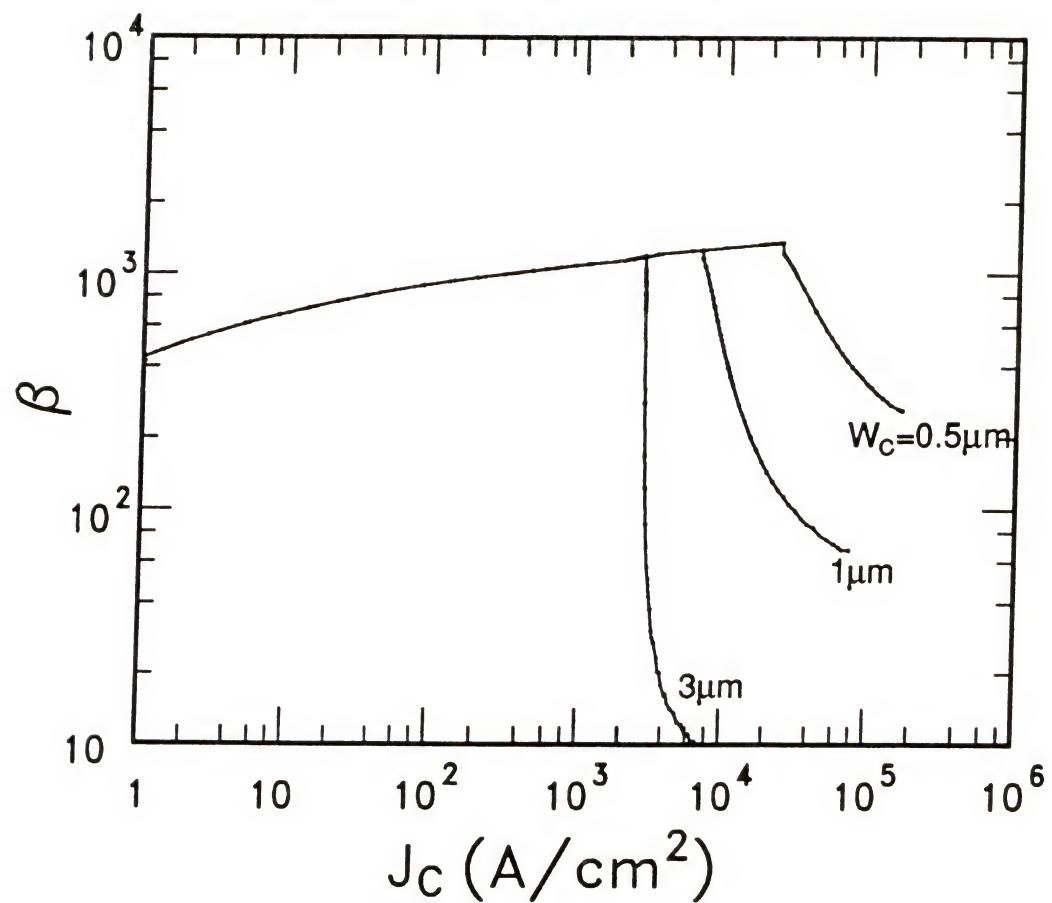
Fig. 7.5 Current gain as a function of the collector current density for three different collector layer thicknesses.

collector, and even for very high forward biases across the metallurgical junction, the collector will not go into high injection. For example, for a p-GaAs ($N_A=10^{19}$ cm^{-3}) / N-Al$_{0.3}$Ga$_{0.7}$As ($N_D=10^{16}$ cm^{-3}) base-collector junction, achieving a hole density of 10^{16} cm^{-3} on the collector side would require such a large forward bias that the electron density on the base side would have to be 10^{23} cm^{-3}; this is physically impossible (the numbers are calculated using the law of the junction, pn = n_i^2 exp(qV_{BC}/kT)). The impossibility of high injection implies that no base pushout would occur in DHBTs. One would then expect the following behavior to occur in DHBTs.

We assume a p-GaAs/N-AlGaAs base/collector junction. We also assume that the v_d-E characteristics of electrons in AlGaAs are qualitatively like those of silicon [Shu87].

Keeping the base-collector reverse bias (V_{BC}) fixed, as the base-emitter bias (V_{BE}) is increased, the collector current (I_C) will increase; hence the voltage drop across, and the electric field in, the quasi-neutral collector epi-layer will increase, and the base-collector junction will become less reverse biased (or even forward biased). With increasing V_{BE}, and thus increasing I_C, a point would be reached where the electric field in the collector will be so high that the electron velocity will saturate. Beyond the point of velocity saturation, the only way for the current to increase will be by an increase in the electron density; and since no substantial hole injection can occur, the presence of excess electrons will cause a space-charge limited current

flow (similar to that in region II of Fig. 7.2. Only now, the entire collector will be like region II).

From the above qualitative description, we propose the following model for current flow through the collector of a DHBT.

If the current flow is ohmic, then as in Eqn. (7.3-3b)

$$J_C = - q \ \mu_n \ N_{DC} \ E = - q \ \mu_n \ N_{DC} \ \frac{V_{CC'}}{W_C - X_{NC}} \tag{7.5-1}$$

where $V_{CC'}$ is defined in Eqn. (7.3-6a), and X_{NC} is the junction space charge region thickness on the collector side.

If the current flow is nonohmic, we assume that the current density can be written as

$$J_C = - q \ n \ v_{sat} \tag{7.5-2}$$

where n is the electron density and v_{sat} is the saturated drift velocity in AlGaAs. Here, we must find n from Poisson's equation. Hence we write

$$\frac{dE}{dx} = \frac{q}{\varepsilon} \ (\ N_{DC} - n \ )$$

$$= \frac{q}{\varepsilon} \left(N_{DC} + \frac{J_C}{q v_{sat}} \right)$$

Integrating yields

$$E = \frac{q}{\varepsilon} \left(N_{DC} + \frac{J_C}{q v_{sat}} \right) (\ x - X_{NC} \ ) - E_{sat} \tag{7.5-3}$$

where it is assumed that the electric field at $x = X_{NC}$ is E_{sat} (just as we assumed that $E = E_p$ at $x = W_{CIB}$ in Sec. 7.2). Integrating Eqn. (7.5-3) yields

$$V(x) = - \int E \ dx = \frac{-q}{2\varepsilon} \left(N_{DC} + \frac{J_C}{q v_{sat}} \right) (\ x - X_{NC} \ )^2 + E_{sat} \ (x - X_{NC})$$

where we have imposed the boundary condition, $V(x=X_{NC}) = 0$. Then from
the definition of $V_{CC'}$,

$$V_{CC'} = V(x=W_C) = -q/2\varepsilon \ (N_{DC}-n)(W_C-X_{NC})^2 + E_{sat}(W_C-X_{NC})$$

Solving this for n and substituting into Eqn. (7.5-2) gives

$$J_C = - q \ v_{sat} \left\{ N_{DC} + \frac{2\varepsilon \ [\ V_{CC'} - E_{sat} \ (W_C-X_{NC}) \ ]}{q \ N_{DC} \ (W_C-X_{NC})^2} \right\} \qquad (7.5-4)$$

Eqns. (7.5-1) and (7.5-4) characterize the current flow through the
collector under ohmic and nonohmic conditions, respectively. These
equations, when coupled with the equations for transport through the
base, yield the terminal characteristics.

We note two points here.

(1) Tiwari [Tiw88] did high current density measurements on DHBTs in
which the base-collector heterojunction was graded. He found that the
collector current saturated at a certain V_{BE}, and the value of the V_{BE}
was a function of V_{BC}. He ascribed this behavior to what he called the
"alloy barrier" effect: the energy gap grading in the collector caused
an energy barrier to appear in the conduction band when the junction
became heavily forward biased (due to the ohmic drop in the collector).
The barrier so created prevented electron injection into the collector
and hence caused the current to saturate. In the analysis that we
presented, we did not consider this effect.

(2) There has been no published date on high current density effects in
abrupt heterojunction DHBTs. Thus it is not clear whether the
description given in this section is what actually happens, or it is
merely a conjecture. The reason, given this uncertainty, for including

this section in the dissertation is so that it could be used as a possible guide in future modeling efforts, once more data becomes available.

If valid, this model will be applicable to Si/SiGe/Si transistors also, which are naturally double heterojunction transistors, and in which the collector, being made of silicon, has the same v_d-E characteristics as we used in the analysis.

CHAPTER 8
NON-QUASI-STATIC MODELS

8.1 Introduction

Conventional bipolar transistor models such as the Gummel-Poon model [Gum70] and the model derived in chapter 5 assume the quasi-static approximation in carrying out the transient analysis. The quasi-static approximation assumes that carriers injected across a junction move with infinite velocity during transients such that the excess carrier density profile has the same voltage dependence as in steady-state at each instant in time. The time, or delay, necessary for carriers to reach that steady-state distribution, is therefore neglected.

Models that avoid this approximation have been proposed for the homojunction bipolar transistor since the early days of transistor analysis [Win59, Fos86, Klo87, Che88, Wu89]. One of the more accurate methodologies [Che88, Wu89] used for deriving such models has been to simultaneously solve the time-dependent continuity equation and the current density equation by using Laplace transformation. In this chapter, we couple this methodology with Lundstrom's concept of the junction velocity ([Lun86], see also chapters 4 and 5) to formulate a non-quasi-static model for the heterojunction bipolar transistor. The analysis is straightforward. The purpose here is to illustrate the feasibility of such modeling. Thus only the simplest case is considered here. Wu and Lindholm have written a series of papers [Wu89, Wu90, Wu91] on non-quasi-static models for homojunction bipolar transistors

152

operating under many different conditions. The methodology presented
here can therefore be used to model more complex operating conditions in
HBTs by coupling it with Wu and Lindholm's procedures.

8.2 NQS Model for an HBT

We derive a non-quasi-static model for the quasi-neutral base
(QNB) region of a single heterojunction bipolar transistor. The
following simplifying assumptions are made in the analysis.

(i) Low injection and uniform doping density in the QNB.

(ii) No base width modulation.

(iii) No recombination in the space charge region or the QNB.

The time-dependent, one-dimensional continuity equation for
electrons, neglecting recombination, is written as

$$\frac{\partial n(x,t)}{\partial t} = \frac{1}{q} \frac{\partial j_N(x,t)}{\partial x} \qquad (8.2\text{-}1)$$

Taking Laplace transforms, and neglecting initial conditions

$$s\, \underline{N}(x,s) = \frac{1}{q} \frac{d\underline{J}_N(x,s)}{dx} \qquad (8.2\text{-}2)$$

where $\underline{N}$ and $\underline{J}_N$ are respectively the Laplace transforms of n and j_N. Thus
for example

$$\underline{N}(x,s) = \int_0^\infty e^{-st}\, n(x,t)\, dt \qquad (8.2\text{-}3)$$

It is assumed in deriving Eqn. (8.2-2) that

$$\underline{L}\left[\frac{\partial j_N(x,t)}{\partial x}\right] = \frac{d\underline{J}_N(x,s)}{dx}$$

where $\underline{L}$ represents Laplace transformation. For low injection and uniform doping density, the electron current in the QNB will be purely diffusive. Thus

$$j_N(x,t) = q \; D_n \; \frac{\partial n(x,t)}{\partial x} \qquad\qquad (8.2\text{-}4)$$

Taking Laplace transforms

$$\underline{J}_N(x,s) = q \; D_n \; \frac{\partial \underline{N}(x,s)}{\partial x} \qquad\qquad (8.2\text{-}5)$$

Combining (8.2-2) and (8.2-5) to eliminate $\underline{J}_N$ gives

$$s \; \underline{N} = D_n \; \frac{d^2\underline{N}}{dx^2}$$

Hence

$$\frac{d^2\underline{N}}{dx^2} - \frac{s}{D_n} \; \underline{N} = 0 \qquad\qquad (8.2\text{-}6)$$

This differential equation is easily solved:

$$\underline{N}(x,s) = A \; e^{x/\lambda} + B \; e^{-x/\lambda} \qquad\qquad (8.2\text{-}7)$$

where

$$1/\lambda^2 = s/D_n \qquad\qquad (8.2\text{-}7a)$$

The boundary conditions at the emitter and collector edges of the SCRs can be written, in low injection, as (see Eqn. (4.4-5))

$$\underline{N}(0,s) = n_o \; \underline{L} \; \{ \; \exp[qv_{BE}(t)/kT] - 1 \; \} + \underline{J}_N(0,s)/qS_{EN} \qquad (8.2\text{-}8a)$$

and

$$\underline{N}(W_B,s) = n_o \; \underline{L} \; \{\exp[qv_{BC}(t)/kT] - 1\} \qquad\qquad (8.2\text{-}8b)$$

where n_0 is the equilibrium electron density in the base and S_{EN} is the junction velocity we discussed in Sec. 4.4 (Eqn. (4.4-4) and the discussion following that equation). From Eqn. (8.2-7) we can get A and B in terms of $\underline{N}(0,s)$ and $\underline{N}(W_B,s)$:

$$A = \frac{- \exp(-W_B/\lambda)\ \underline{N}(0,s)\ +\ \underline{N}(W_B,s)}{2\ \sinh(W_B/\lambda)} \qquad (8.2\text{-}9a)$$

and

$$B = \frac{\exp(W_B/\lambda)\ \underline{N}(0,s)\ -\ \underline{N}(W_B,s)}{2\ \sinh(W_B/\lambda)} \qquad (8.2\text{-}9b)$$

Combining Eqns. (8.2-5), (8.2-7) and (8.2-9) we get

$$\underline{J}_N(0,s) = qD_n/\lambda\ (\ A\ -\ B\ )$$

$$= qD_n/\lambda\ [\ -\coth(W_B/\lambda)\ \underline{N}(0,s)\ +\ csch(W_B/\lambda)\ \underline{N}(W_B,s)\ ]$$

$$(8.2\text{-}10)$$

Now combining Eqns. (8.2-8) and (8.2-10) and rearranging gives

$$\underline{J}_N(0,s) = \frac{qD_n n_0/\lambda\ \{\ -\coth(W_B/\lambda)\ \underline{L}[\exp(qv_{BE}/kT)-1]\ +\ +\ csch(W_B/\lambda)\ \underline{L}[\exp(qV_{BC}/kT)-1]\ \}}{[\ 1\ +\ D_n/(S_{EN}\lambda)\ \coth(W_B/\lambda)\ ]}$$

$$(8.2\text{-}11a)$$

$$\overset{\Delta}{=} -\ \frac{qD_n n_0}{L_{EE}}\ \underline{L}[\exp(qv_{BE}/kT)-1]\ +\ \frac{qD_n n_0}{L_{EC}}\ \underline{L}[\exp(qv_{BC}/kT)-1] \quad (8.2\text{-}11b)$$

where

$$L_{EE} = \frac{\lambda}{\coth(W_B/\lambda)} + \frac{D_n}{S_{EN}} \qquad L_{EC} = \frac{\lambda}{csch(W_B/\lambda)} + \frac{D_n}{S_{EN}}\ \cosh(WB/\lambda)$$

$$(8.2\text{-}11c)$$

Eqn. (8.2-11a) is analogous to the model derived by Chen et al [Che87] for the homojunction transistor under the same assumptions we made here. The only difference is the denominator of Eqn. (8.2-11a) which in a homojunction transistor will simply be unity. The additional term here arises from the boundary condition of Eqn. (8.2-8a) which is peculiar to a heterojunction. We discussed a similar analogy in chapter 5.

We now use the series expansions for the hyperbolic functions, and truncate these in order to derive a circuit model based on Eqn. (8.2-11).

$$\sinh(x) \simeq x + x^3/3! + ...$$
$$\cosh(x) \simeq 1 + x^2/2! + ...$$

Then from Eqn. (8.2-11c)

$$L_{EE} = \frac{\lambda \sinh(W_B/\lambda)}{\cosh(W_B/\lambda)} + \frac{D_n}{S_{EN}} \simeq \frac{\lambda(W_B/\lambda + W_B^2/6\lambda^3)}{(1+W_B^2/2\lambda^2)} + \frac{D_n}{S_{EN}}$$

Hence

$$\frac{1}{L_{EE}} = \frac{1}{W_B} \frac{(1 + W_B^2/2\lambda^2)}{(1+D_n/W_B S_{EN}) + W_B^2/6\lambda^2 (1+3D_n/W_B S_{EN})} \qquad (8.2-12a)$$

Similarly

$$\frac{1}{L_{EC}} = \frac{1}{W_B} \frac{1}{(1+D_n/W_B S_{EN}) + W_B^2/6\lambda^2 (1+3D_n/W_B S_{EN})} \qquad (8.2-12b)$$

Eqns. (8.2-5,7,9,11) also give the current at the collector edge as

$$\underline{J}_N(W_B,s) \overset{\Delta}{=} - \frac{q D_n n_o}{L_{CE}} \underline{L}[\exp(q v_{BE}/kT)-1] + \frac{q D_n n_o}{L_{CC}} \underline{L}[\exp(q v_{BC}/kT)-1] \quad (8.2-13)$$

where

$$\frac{1}{L_{CC}} = \frac{\coth(W_B/\lambda) + D_n/S_{EN}\lambda}{\lambda \ [1+D_n/S_{EN}\lambda\coth(W_B/\lambda)]} \qquad \text{and} \qquad L_{CE} = L_{EC} \qquad (8.2\text{-}13a)$$

We now rewrite Eqns. (8.2-11b) and (8.2-13) as

$$I_E = \underline{J}_N(0,s) \ A$$

$$= -I_0 \ (W_B/L_{EE}) \ \underline{L}[\exp(qv_{BE}/kT)-1] + I_0 \ (W_B/L_{EC}) \ \underline{L}[\exp(qv_{BC}/kT)-1]$$

$$= - \ I_{11} + I_{12} \qquad\qquad (8.2\text{-}14a)$$

$$I_C = \underline{J}_N(W_B,s) \ A$$

$$= -I_0 \ (W_B/L_{CE}) \ \underline{L}[\exp(qv_{BE}/kT)-1] + I_0 \ (W_B/L_{CC}) \ \underline{L}[\exp(qv_{BC}/kT)-1]$$

$$= - \ I_{21} + I_{22} \qquad\qquad (8.2\text{-}14b)$$

where A is the cross-sectional area, and $I_0 = qD_n n_o A/W_B$. An equivalent circuit for this set of equations is shown in Fig. 8.1(a) [Wu89]. The values of the current sources are

$$i_{11} = - \ I_{11} + I_{12} = I_0 W_B \ (-1/L_{EE}+L_{EC}) \ \underline{L}[\exp(qv_{BE}/kT)-1] \qquad (8.2\text{-}15a)$$

$$i_{12} = I_{12} - I_{21} = I_0 W_B/L_{EC} \ \{\underline{L}[\exp(qv_{BE}/kT)-1]-\underline{L}[\exp(qv_{BC}/kT)-1]\}$$

$$(8.2\text{-}15b)$$

$$i_{22} = - \ I_{21} + I_{22} = I_0 W_B \ (-1/L_{CE}+L_{CC}) \ \underline{L}[\exp(qv_{BC}/kT)-1] \qquad (8.2\text{-}15c)$$

We now derive an equivalent circuit in the time domain. From Eqns. (8.2-7a,12a,12b) we can write

$$W_B \left(- \frac{1}{L_{EE}} + \frac{1}{L_{EC}} \right) = \frac{1}{1 + D_n/W_B S_{EN}} \ \frac{s\tau_{11}}{1 + s\tau_{12}} \qquad (8.2\text{-}16)$$

where

$$\tau_{11} = \frac{W_B{}^2}{2D_n} \ , \qquad\qquad \tau_{12} = \frac{W_B{}^2}{6D_n} \ \frac{(1+3D_n/W_B S_{EN})}{(1+D_n/W_B S_{EN})} \qquad (8.2\text{-}17)$$

Substituting in Eqn. (8.2-15a) gives

$$i_{11} = \frac{I_0}{(1+D_n/W_B S_{EN})} \frac{s\tau_{11}}{(1+s\tau_{12})} \underline{L}[\exp(qv_{BE}/kT)-1]$$

Hence

$$\frac{1}{I_0'} \left[\frac{1}{s\tau_{11}} + \frac{\tau_{12}}{\tau_{11}} \right] i_{11}(s) = \underline{L}[\exp(qv_{BE}/kT)-1]$$

where $I_0' = I_0/(1+D_n/W_B S_{EN})$. Taking inverse Laplace transforms

$$\frac{1}{I_0'\tau_{11}} \int_0^t i_{11}(t') \, dt' + \frac{\tau_{12}}{I_0'\tau_{11}} i_{11}(t) = \frac{[\exp(qv_{BE}/kT)-1]}{v_{BE}} v_{BE}$$

Hence we can write

$$\frac{1}{C_{11}} \int_0^t i_{11}(t') \, dt' + R_{11} \, i_{11}(t) = v_{BE}(t)$$

where

$$C_{11} = I_0'/v_{BE} \, \tau_{11} \, [\exp(qv_{BE}/kT)-1]$$

and

$$R_{11} = \tau_{12} / C_{11}$$

The current sources i_{12} and i_{22} can be replaced similarly by circuit elements. The complete equivalent circuit for the quasi-neutral base is shown in Fig. 8.1(b). The expressions for the various circuit elements are

$$R = \frac{(v_{BE} - v_{BC})}{I_0' \, [\exp(qv_{BE}/kT) - \exp(qv_{BC}/kT)]}$$

$$L = \tau_{12} \, R$$

$$C_{22} = I_0'/v_{BC} \, \tau_{22} \, [\exp(qv_{BE}/kT) - 1]$$

$$R_{22} = \tau_{12} / C_{22}$$
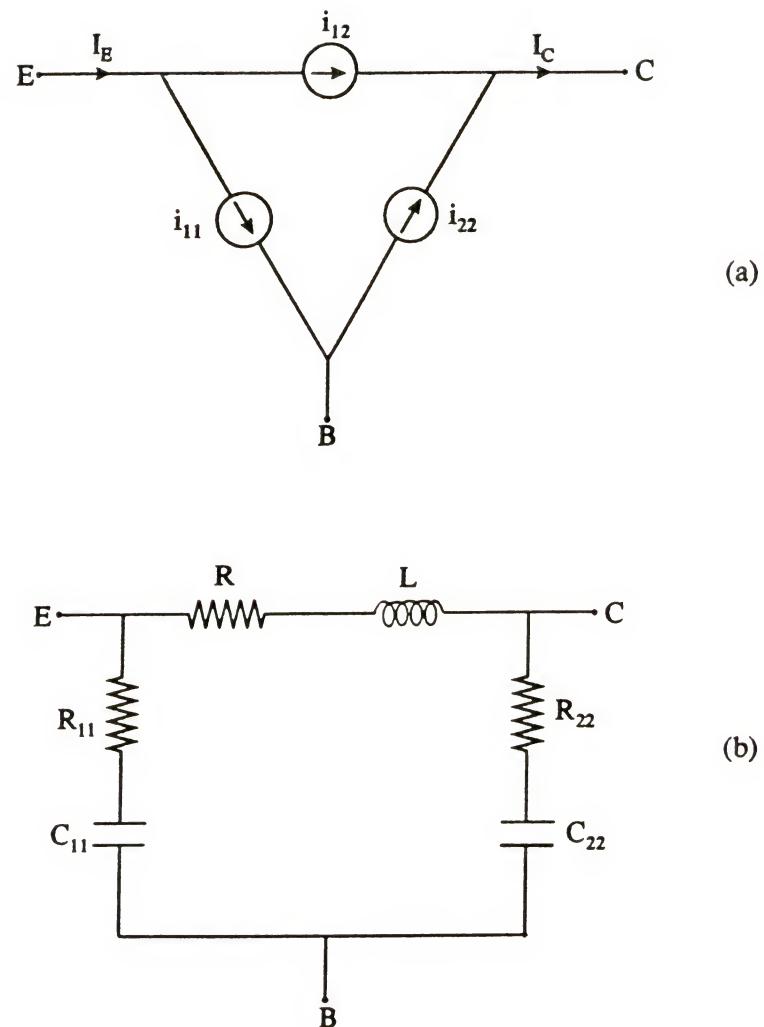
Fig. 8.1 A non-quasi-static equivalent circuit with (a) current sources and (b) passive circuit elements, for the base of an HBT.

$$\tau_{22} = \frac{W_B^2}{2D_n} \left(1 + \frac{2D_n}{W_B S_{EN}} \right)$$

As we said at the beginning of the chapter, the analysis carried out above is illustrative in its aim. More complex cases, such as high injection and/or nonuniform doping density in the base, double heterojunction transistor, base pushout, etc., can be analyzed by coupling the method presented above with the methodologies proposed by Wu and Lindholm [Wu89, Wu90, Wu91].

CHAPTER 9
SOME PROBLEMS ADDRESSED BUT NOT SOLVED

In this chapter we look at three issues which were addressed during the course of this dissertation but which have not yielded any definite results.

9.1 Fabrication of Heterojunctions

In order to corroborate the analytical models derived in the earlier chapters, it was desirable that experimental data be obtained. Given the limited facilities and expertise in compound semiconductor fabrication technology, it was only possible to fabricate heterojunction diodes. Thus the task of fabricating N-AlGaAs/p-GaAs heterojunctions was undertaken.

An MBE-grown structure, as shown in Fig. 9.1, was fabricated (by Dr. Park's group in the Department of Materials Science and Engineering). Both the p and the n side of the junction were doped to 10^{17} /cm^3. The top n$^+$ layer facilitated low resistance ohmic contacts. The ohmic contact on the N$^+$ layer was made by depositing AuGe(88% Au-12% Ge) (1000 A°)/ Ni(500 A°)/ Au(2000 A°) by electron beam evaporation. On the p-type substrate a Cr(200 A°)/ Au(2000 A°) contact was made. A mask was prepared with different sized circular dots, and the pattern was created on the N-side of the wafer by standard photolithography. The metals were etched using a KI-I$_2$ solution (400 gms KI + 100 gms I$_2$ in 400 ml of H$_2$O). GaAs and AlGaAs were etched using NH$_4$OH:H$_2$O$_2$:H$_2$O (3:1:15), to obtain mesa diodes. A Dektak surface profiler was used to

Au / Ni / AuGe

0.4 μm	n^+-GaAs (Si)	2.3×10^{18} /cm^3
0.5 μm	n-Al$_{0.3}$Ga$_{0.7}$As (Si)	10^{17} /cm^3
0.5 μm	p-GaAs (Be)	10^{17} /cm^3
0.5 μm	p^+-GaAs buffer (Be)	10^{19} /cm^3

p-GaAs substrate 10^{18} /cm^3

Au / Cr
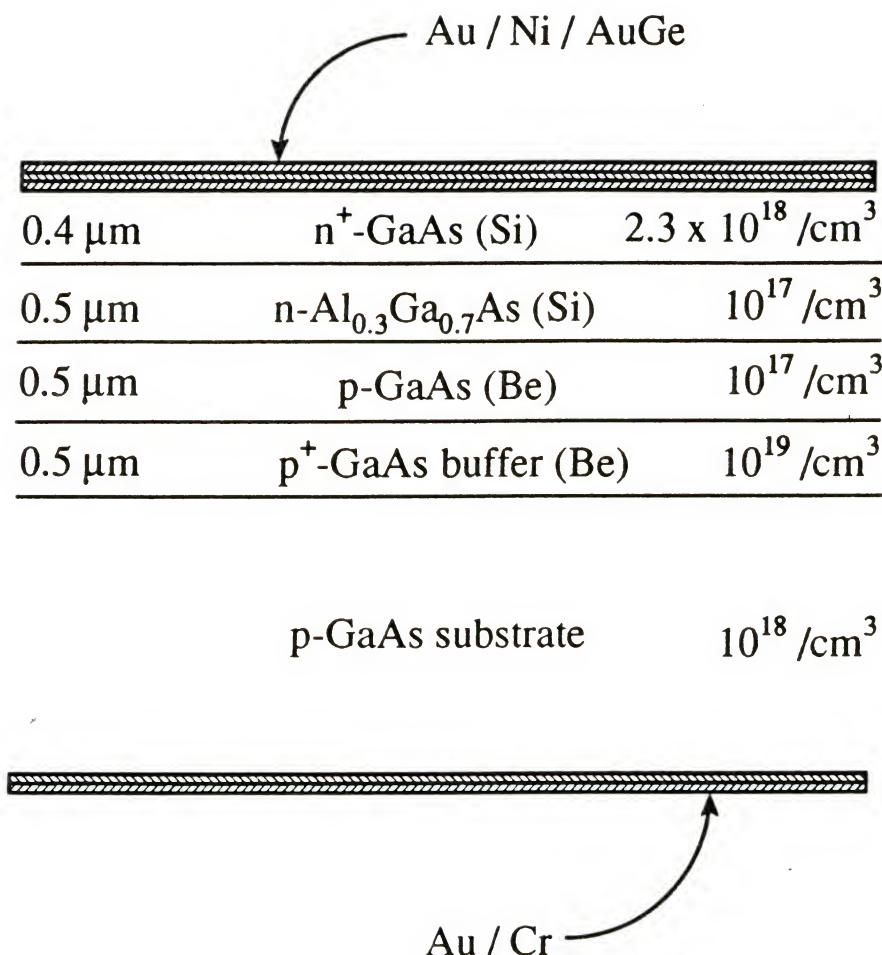
Fig. 9.1 The MBE-grown structure showing the epitaxial layer parameters.

ensure that the correct depth was etched. Fig. 9.2 shows a typical
surface profile, with about 1.5 µm etched.

After identifying the working diodes, I-V characteristics were
measured on an HP 4145. A typical forward log(I)-V curve is shown in
Fig. 9.3. It is seen to be a straight line, indicating a $I = I_0 exp(qV/nkT)$ behavior. The value of n on all diodes was around 2. This
implied that the dominant current was either due to space charge region
recombination or due to surface recombination. Measurements on diodes of
various areas showed that the saturation current I_0 varied linearly with
the radius of the circular dot (and not with the square of the radius).
This indicated that the current must be due to surface recombination.

Now it is known [Hen77, Log77, Henr78a, Henr78b, Hay87, Lyo89a,
LeeW89] that even in highly sophisticated fabrication facilities, GaAs
surfaces close to the active p/n junction cause a significant
recombination current; and there is no established method of obtaining
low surface recombination velocities on GaAs surfaces (although at least
two different approaches have been proposed [Tiwa87, San87, Nott88,
Ma89]). Given this inability to get rid of surface recombination, it was
not possible to obtain any meaningful data from the measurements, except
perhaps by extensive and prolonged measurement procedures. One
alternative was to research passivation of GaAs surfaces, which would
perhaps be a dissertation by itself. Another alternative was to use the
diodes fabricated and do extensive measurements - on various sized
diodes, at various temperatures - and separate out the surface
recombination currents so that information about some bulk properties,
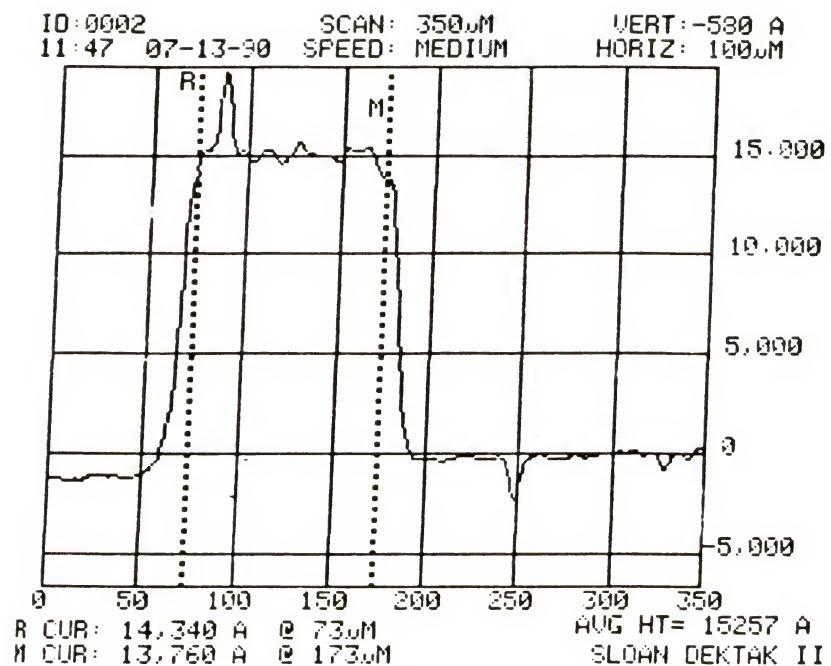
Fig. 9.2 Surface profile of a mesa diode made from the epitaxial
structure of Fig. 9.1. The profile was obtained on a Sloan
Dektak surface profiler.

****** GRAPHICS PLOT ******
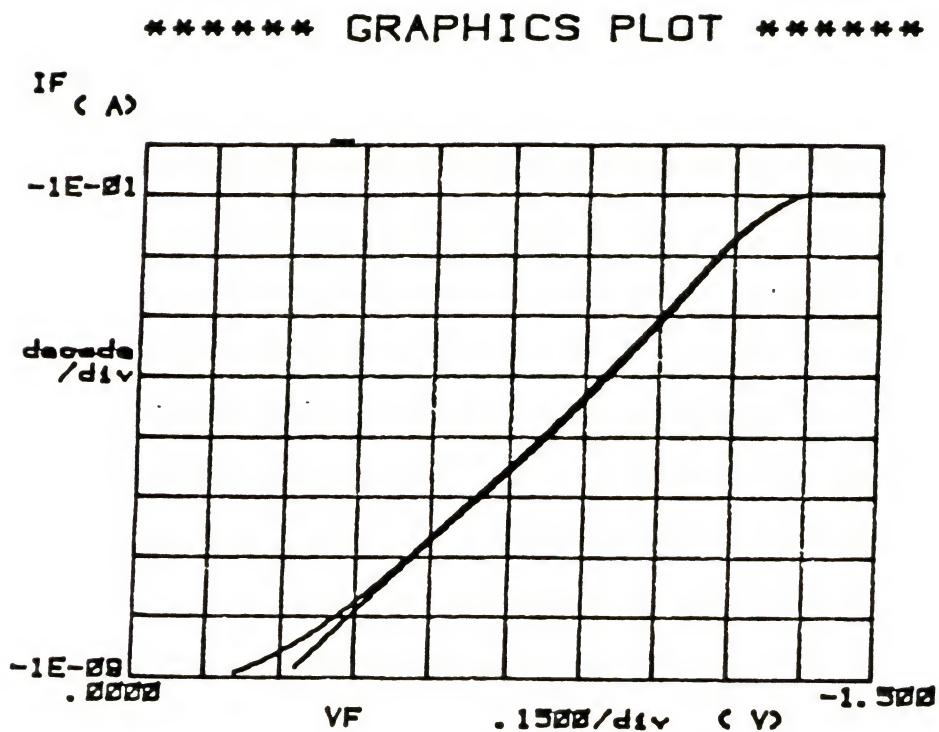
Fig. 9.3 Typical log(I)-V characteristics of the heterojunction.

such as carrier lifetime, effect of heterojunction on recombination, etc., could be obtained. BUT, the whole inclination and interest of this author have always been in pursuing physical-analytical research; experimental work has never been either his forte or his interest. Thus, after spending sufficient time on the project to realize that this lay beyond the capabilities of the author, it was given up as being unfeasible.

9.2 Capacitance of p/n Junction Space-Charge Regions

While working on the issue of base pushout (chapter 7) and its effects on transistor performance, the question of how it would affect the cut-off frequency, f_T, came up. To evaluate f_T, it is necessary to have models for the various capacitances of the transistor, especially the space-charge region (SCR) capacitances. Thus a search of the literature for models of homojunction and heterojunction SCR capacitances was done, especially under large forward bias conditions, since the base-collector junction is heavily forward biased when base pushout occurs (see chapter 7).

Although many models have been proposed, and many different approaches used, to determine the SCR (or the junction) capacitance, two major inadequacies were found in the literature. These were as follows: (1) Many different definitions of junction capacitance were found, and there was no consensus on what the correct definition is. In fact, this author has not found much discussion of the definition itself in the literature.

(2) Little attention has been given to modeling the SCR capacitance at moderate or high forward biases. Thus most models proposed have simply been modifications of the basic depletion capacitance derived assuming the depletion approximation (which becomes less and less valid as the forward bias is increased, since then there are more and more free carriers in the SCR).

In this section therefore, we will first enumerate some of the many definitions of the SCR capacitance used in the literature, and then propose a new definition that this author has come up with.

As mentioned at the beginning of this chapter, modeling the SCR capacitance was one of the issues that was addressed during the course of this work, but not solved. We will discuss the difficulties involved in such modeling at the end of this section.

9.2.1 Definitions of the SCR Capacitance: A Literature Review

From field theory, capacitance is the rate of change of the stored charge with respect to the applied voltage. Thus

$$C = dQ/dV$$

For a p/n junction, V is the voltage applied at the terminals. It is Q that needs a more specific definition. We shall use the terminology defined in Fig. 9.4 in the following.
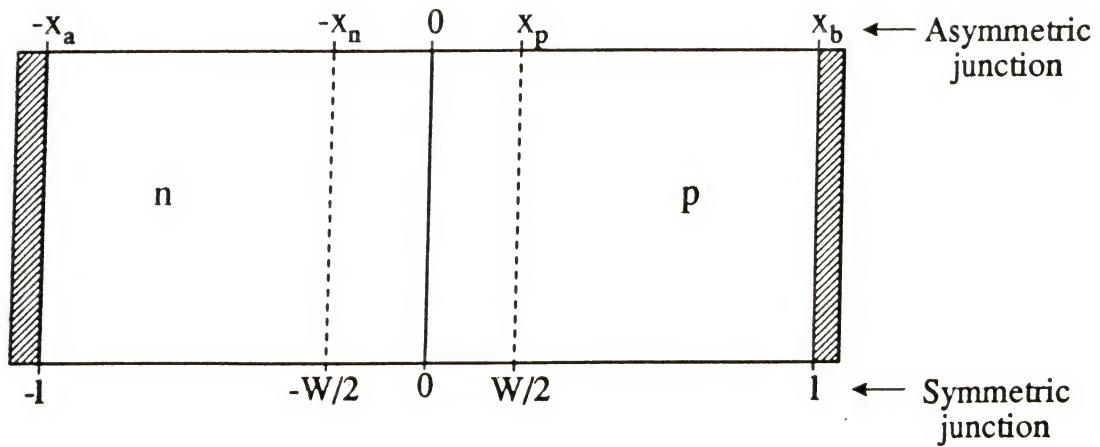
Fig. 9.4 A schematic of an n/p junction showing the various geometrical parameters.

(i) Shockley [Sho49] defined the p/n diode capacitance as

$$C = q \frac{d}{dV} \int_{-x_a}^{x_b} p \, dx \qquad (9.2-1)$$

Then to calculate its value, he assumed the depletion approximation, and wrote

$$C_{SCR} = q \frac{d}{dV} \int_0^{x_p} p \, dx = q \, N_A \frac{dx_p}{dV}$$

where N_A is the acceptor concentration. The expression is therefore valid only under reverse bias conditions, when there are no free carriers in the SCR.

(ii) Sah [Sah61b] considered a linearly graded, symmetric junction. He defined the total capacitance of the diode as

$$C = q \frac{d}{dV} \int_{-1}^{1} p \, dx$$

which is the same as Shockley's definition. He then defined the SCR capacitance as

$$C_{SCR} = q \frac{d}{dV} \int_{-W/2}^{W/2} p \, dx \qquad (9.2-2)$$

This is the definition used by Lindholm and Liou [Lin88] also. We shall discuss this definition a little more later.

(iii) Gummel and Scharfetter [Gumm67] defined

$$C_{SCR} = q \frac{d}{dV} \int_{-\infty}^{0} (N_D - n) \, dx$$

which is the same as saying

$$C_{SCR} = q \frac{d}{dV} \int_0^{-x_a} n \, dx \qquad (9.2-3)$$

since $\int_{0}^{-x_a} N_D \, dx$ does not change with voltage. Note that Eqn. (9.2-3) is

not equal in value to Eqn. (9.2-2) since only one side of the SCR is integrated over in the former. In practice though, for reverse and small forward biases, they would be equal as there are few electrons on the p-side of the SCR.

(iv) Chang [ChangY67] used

$$C_{SCR} = q \frac{d}{dV} \int_{-x_n}^{x_p} [\, n(x) - n_0(x) \,] \, dx \qquad (9.2-4)$$

where $n_0(x)$ is the electron density in equilibrium. Once again, this definition would give a different result from the earlier ones, because although n_0 is not a function of V, x_n and x_p are, so that the n_0 term will contribute to the total capacitance.

(v) Nuyts and Overstraeten [Nuy71] used a very enigmatic definition:

$$C_{SCR} = q \frac{d}{dV} \int_{x_0}^{\infty} (p-n) \, dx \qquad (9.2-5)$$

where x_0 is the point within the SCR where p=n. This author has not been able to find the rationale of this definition.

(vi) van den Biesen [Bie85a, Bie85b] has still another definition:

$$C_{SCR} = \quad C_{depl} \quad + \quad C_N$$

$$= q \frac{d}{dV} \left(\int_{x_1}^{x_b} (p-n) \, dx + \int_{-x_n}^{x_1} p \, dx + \int_{x_1}^{x_p} n \, dx \right) \qquad (9.2-6a)$$

where x_1 is the point at which $\partial p/\partial V = \partial n/\partial V$. This can be rewritten as

$$C_{SCR} = q \frac{d}{dV} \left(\int_{-x_n}^{x_p} p \, dx + \int_{x_p}^{x_b} (p-n) \, dx \right) \qquad (9.2-6b)$$

Here again, the second integral makes this definition different from the earlier ones; and no discussion is provided for using such a definition.

(vii) Finally, there is Lindholm and Liou's definition [Lin88]:

$$C_{SCR} = \quad C_D \quad + \quad C_F$$

$$= q \int_{-x_n}^{x_p} \partial p/\partial V \, dx + q \, p(x_p) \frac{dx_p}{dV} \qquad (9.2\text{-}7)$$

The second term in the above definition has a negative value, since as V increases, x_p decreases. We discuss this more below.

9.2.2 A new definition for the SCR Capacitance

We propose yet another definition for the SCR capacitance:

$$C_{SCR} = q \int_{-x_n}^{x_p} \frac{\partial p}{\partial V} \, dx \qquad (9.2\text{-}8)$$

The method of arriving at this definition is as follows. The total capacitance of the p/n diode, that is the space-charge plus the diffusion capacitance, is the rate of change of the stored charge with respect to the applied voltage. Since there are two types of carriers involved (electrons and holes), and since the total system remains charge-neutral, the total diode capacitance can be written as the rate of change of the number of electrons, or holes, within the diode. Thus

$$C_{total} = q \frac{d}{dV} \int_{-x_a}^{x_b} (n \text{ or } p) \, dx \qquad (9.2\text{-}9)$$

where x_a and x_b represent the ohmic contacts, as indicated in Fig. 9.4. Now Leibnitz' rule allows us to write

$$C_{total} = q \int_{-x_a}^{x_b} \partial p/\partial V \, dx \qquad (9.2\text{-}10)$$

since x_a and x_b are not functions of x; and we use the hole density p for concreteness (using n will be equally valid). The integral is now split into three parts:

$$C_{total} = q \int_{-x_a}^{-x_n} \partial p/\partial V \, dx + q \int_{-x_n}^{x_p} \partial p/\partial V \, dx + q \int_{x_p}^{x_b} \partial p/\partial V \, dx \qquad (9.2\text{-}11)$$

The claim now is that the first and the third integrals above represent the change in the stored charge in the quasi-neutral regions, and thus correspond to the diffusion capacitance, while the second integral represents the change in the stored charge in the space charge region (SCR) and thus defines the SCR capacitance. Hence the definition of Eqn. (9.2-8).

Let us now analyze Eqn. (9.2-9) a little differently. The integral in that equation can be split into three parts, such that

$$C_{total} = q \frac{d}{dV} \int_{-x_a}^{-x_n} p \, dx + q \frac{d}{dV} \int_{-x_n}^{x_p} p \, dx + q \frac{d}{dV} \int_{x_p}^{x_b} p \, dx \qquad (9.2\text{-}12)$$

The definition of (ii) and (vii) use the second term in the above expression as the definition of C_{SCR}. They thus say

$$C_{SCR} = q \frac{d}{dV} \int_{-x_n}^{x_p} p \, dx \qquad (9.2\text{-}13)$$

The difference between Eqn.s (9.2-8) and (9.2-13) is that in the latter, the derivative is outside the integral, while in the former, it is inside. Use of Leibnitz' rule on Eqn. (9.2-13), as Lindholm and Liou did [Lin88], gives

$$C_{SCR} = q \int_{-x_n}^{x_p} \partial p/\partial V \, dx + q \, p(x_p) \, dx_p/dV + q \, p(-x_n) \, dx_n/dV \quad (9.2\text{-}14)$$

Such a definition has two additional terms compared to the definition of Eqn. (9.2-8). But these two terms, it seems to this author, are superfluous, and are merely a result of the mathematical construct used to write Eqn. (9.2-12). For example, using the Leibnitz' rule on the third term in Eqn. (9.2-12) yields

$$q \, \frac{d}{dV} \int_{x_p}^{x_b} p \, dx = q \int_{x_p}^{x_b} \partial p/\partial V \, dx - q \, p(x_p) \, dx_p/dV$$

The second term on the right hand side above is equal and opposite to the second term on the right hand side of Eqn. (9.2-14). In the total capacitance therefore, these terms do not appear at all, as is apparent from Eqn. (9.2-11) also.

We end this section with a comment about why modeling C_{SCR} is a non-trivial task. Rewriting the new definition of C_{SCR} from Eqn. (9.2-8)

$$C_{SCR} = q \int_{-x_n}^{x_p} \frac{\partial p}{\partial V} \, dx \quad (9.2\text{-}8)$$

p can be written as

$$p = n_i \exp [\, (E_i - E_{FP})/kT \,]$$

For a uniform energy gap material, n_i will be constant, and E_i could be written as $E_i = -qv$, where v is the electrostatic potential. Hence

$$\partial p/\partial V = p/kT [\, -q\partial v/\partial V - \partial E_{FP}/\partial V \,]$$

v is a function of both x and V, and could determined only by solving Poisson's equation. But Poisson's equation, within the SCR, is solvable

only if the depletion approximation is assumed. Otherwise it can be integrated only once to yield dv/dx, but not any further. Hence it is extremely difficult to evaluate $\partial v/\partial V$ analytically under conditions when the depletion approximation cannot be used (that is, at moderate or large forward biases). This, it seems to this author, is the major hurdle in modeling C_{SCR}.

9.3 Understanding and Modeling Velocity Overshoot

In Sec. 4.7 we reviewed papers that studied velocity overshoot and its effects on electron transport in the base-collector space charge region (SCR) of GaAs HBT's. Perhaps more because of this author's fascination with this phenomenon, rather than any real need to address it, velocity overshoot was studied during the course of this work. As we shall see though, apart from gaining a qualitative understanding of velocity overshoot, no other result was achieved.

In this section we will first describe qualitatively the reasons for velocity overshoot, and then look at the attempts at simplistic analytical modeling of overshoot.

9.3.1 Velocity Overshoot: A Qualitative Look

The following explanation is largely taken from the paper by Maloney and Frey [Malo77]. Other places where a qualitative description can be found are [Gru80, Tei83, Con85].

There are two different reasons why overshoot is observed in semiconductors. We will look at both of these briefly.

Semiconductors such as Si and Ge, have the following characteristic properties.

(i) The most significant scattering mechanism, at room temperature, in these semiconductors is due to acoustic phonons; a scattering event with a phonon randomizes the electron momentum. But these acoustic phonons have energies of around 0.05 eV, so that high energy electrons require several scattering events to randomize their energy. Thus the energy relaxation time in such semiconductors is larger than the momentum relaxation time.

(ii) The momentum relaxation time is a decreasing function of the energy of electrons.

Although qualitative, our analysis is best done with the help of two phenomenological conservation equations for electron motion [Shu76]:

$$\frac{d(m^*v)}{dt} = qE - \frac{m^*v}{\tau_p(\xi)} \tag{9.3-1}$$

$$\frac{d\xi}{dt} = qEv - \frac{\xi-\xi_0}{\tau_\xi(\xi)} \tag{9.3-2}$$

where m^* is the electron effective mass, v is the drift velocity, ξ is the electron energy, E is the electric field, τ_p and τ_ξ are the momentum and energy relaxation times, respectively, and ξ_0 is the thermal equilibrium value of ξ. Note that the electron energy consists of two components: the thermal energy and the drift energy, so that $\xi = 3/2kT_e + 1/2m^*v^2$, where T_e is the electron temperature. Equations (9.3-1) and (9.3-2) are respectively the momentum and energy conservation equations.

Let us, for the purpose of illustration, assume that for all energies, $\tau_p(\xi) \ll \tau_\xi(\xi)$. This would allow us to assume that the time

scales involved in momentum change are much smaller that those involved in energy change.

Let us now consider an n-type bulk semiconductor which is suddenly subjected to a large electric field. On such application, the electrons will accelerate and rapidly gain velocity. Within a few collisions, the momentum will relax and reach a quasi-steady-state with the lattice (i.e. $dv/dt \simeq 0$). Now since $\tau_p \ll \tau_\xi$, the time required to reach this state will be so small that the energy will have changed little. Thus from Eqn. (9.3-1)

$$v_p \simeq \frac{qE\tau_p(\xi_0)}{m^*} \qquad (9.3-3)$$

In our simplistic analysis, v_p will be the peak velocity that the electrons will achieve, almost instantaneously (with respect to the time scales involved in energy change). Then as the electrons gradually gain energy due to collisions, $\tau_p(\xi)$ will decrease and thus v will also decrease, until steady-state is reached. Thus in reaching steady-state, electrons would have first overshot to the velocity given by Eqn. (9.3-3), and then gradually decreased to the steady-state velocity, v_{ss}, given by

$$v_{ss} = \frac{qE\tau_p(\xi_{ss})}{m^*}$$

Note that no overshoot will occur if $\tau_p(\xi) \simeq \tau_\xi(\xi)$, because then, momentum and energy will increase concurrently.

Whether velocity overshoot will occur in any particular situation will depend on the magnitude of the applied electric field, the

disparity between τ_p and τ_ξ, and the energy dependence of the relaxation times.

The above brief description of velocity overshoot is applicable only to semiconductors such as Si and Ge, in which the dominant scattering mechanism is due to acoustic phonons. In compound semiconductors such as GaAs and InP, the momentum and energy relaxation times in the lowermost conduction band valley (Γ) are almost equal. This is because the dominant scattering mechanism there is attributed to polar optical phonons, which have a high energy. In such semiconductors, if the electrons stay in the Γ valley, no overshoot will occur. Overshoot in these semiconductors occurs only when the electrons have enough energy to get transferred to the higher valleys. The two characteristics that cause overshoot are the following.

(i) The relaxation time associated with intervalley scattering is large compared to the relaxation time for scattering within the Γ valley.

(ii) Electrons have a much larger effective mass (and hence a lower mobility) in the higher (L and X) valleys than in the lower Γ valley.

The dynamics of overshoot is thus as follows. When suddenly subjected to a high electric field, electrons accelerate, acquire momentum and energy concurrently, and maintain a quasi-steady-state with the lattice, until their energy exceeds the threshold for intervalley transfer (about 0.35 eV in GaAs). Beyond this point, electrons keep gaining momentum, and experience velocity overshoot, since transfer to higher valleys is slow to occur. As more and more electrons get transferred to the higher valleys, the velocity starts reducing (due to

the large effective mass in the higher valleys) until steady-state is attained. Note that if the voltage drop across the high field region is less than the threshold for intervalley transfer (0.35 V in GaAs), then no overshoot will occur.

9.3.2 Attempts at Analytical Modeling

The starting equations for the analysis are the momentum and energy conservation equations [Blo70, Sno85] derivable by taking first and second moments of the Boltzmann transport equation. These equations are as follows [Sno85]:

Momentum conservation:

$$\frac{\partial(m^\star \mathbf{v})}{\partial t} + m^\star \mathbf{v} \cdot \nabla \mathbf{v} = q\mathbf{E} - 1/n\ \nabla(nkT_e) + m^\star \left(\frac{\partial \mathbf{v}}{\partial t} \right)_{coll}$$

Energy conservation:

$$\frac{\partial \xi}{\partial t} + \mathbf{v} \cdot \nabla \xi = q\ \mathbf{v} \cdot \mathbf{E} - 1/n\ \nabla \cdot (nkT_e \mathbf{v}) + m^\star \left(\frac{\partial \xi}{\partial t} \right)_{coll}$$

where $\mathbf{E}$ is the electric field, $\mathbf{v}$ is the drift velocity, ξ is the electron energy, $m^\star$ is the effective mass, T_e is the electron temperature, and "coll" represents collision. It was perhaps Shur who first proposed [Shu76] that the above two equations could be simplified using phenomenological arguments, and written in one dimension as follows:

Momentum conservation:

$$\frac{\partial(m^\star v)}{\partial t} + m^\star v\ \frac{\partial v}{\partial x} = qE - \frac{m^\star v}{\tau_p(\xi)} \tag{9.3-4}$$

Energy conservation:

$$\frac{\partial \xi}{\partial t} + v \frac{\partial \xi}{\partial x} = qvE - \frac{\xi - \xi_0}{\tau_\xi(\xi)} \qquad (9.3\text{-}5)$$

where τ_p and τ_ξ are respectively the momentum and energy relaxation times, and ξ_0 is the equilibrium value of the electron energy ξ. On the left hand side of both these equations there are two terms, one representing a temporal variation and the other a spatial variation. Thus in the presence of only a spatial gradient in the electric field, the equations will further simplify to

$$m^\star v \frac{\partial v}{\partial x} = qE - \frac{m^\star v}{\tau_p(\xi)} \qquad (9.3\text{-}6)$$

and

$$v \frac{\partial \xi}{\partial x} = qvE - \frac{\xi - \xi_0}{\tau_\xi(\xi)} \qquad (9.3\text{-}7)$$

For solving these two coupled differential equations, the energy dependence of τ_p and τ_ξ must be known. Shur [Shu76] suggested that these dependences could be determined from the steady-state v-E and ξ-E characteristics when there are no gradients in the electric field. In such a situation, the left hand sides of Eqns. (9.3-4) and (9.3-5) would be zero and the equations could be rewritten as

$$\tau_p(\xi) = \frac{v_{ss}}{qE_{ss}} \qquad (9.3\text{-}8)$$

$$\tau_\xi(\xi) = \frac{\xi - \xi_0}{qv_{ss}E_{ss}} \qquad (9.3\text{-}9)$$

where the subscript ss indicates steady-state. Note that v_{ss} and E_{ss} are treated as functions of energy, and are found from Monte Carlo

simulations. Shur's results showed that solutions obtained from Eqns. (9.3-6,7,8,9) were in good agreement with exact Monte Carlo simulation results, in presence of sharp gradients in the electric field.

Now the attempt of this author was to solve Eqns. (9.3-6) and (9.3-7) using numerical techniques. For this it was necessary to know τ_p and τ_ξ as functions of energy. Here there were two possible approaches. One was to develop one's own steady-state Monte Carlo routine and use Eqns. (9.3-8) and (9.3-9). But developing such a program is an extremely cumbersome and time consuming task; and given the relative insignificance of modeling overshoot in the overall goal of modeling heterojunction bipolar transistors, this option was eliminated as being impractical. The other approach was to find data published in the literature. To this end, the literature was combed painstakingly to gather all the papers that had relevant data. But at the end of the search, less than five papers were found on this subject; and the data presented there were far from sufficient to attempt to solve the equations for overshoot. Thus in the end, this whole endeavor was given up as being unfeasible. Perhaps in the future, once more numerical and empirical data become available, this problem would be easier to solve.

Before ending this section, we mention one other approach to model velocity overshoot proposed by Thornber [Thor82]. He suggested the following equation:

$$J(x,t) = q\, n(x,t)\, [\; v(E) + W(E)\, \partial E/\partial x + B(E)\, \partial E/\partial t \;]$$
$$- D(E)\, \partial n/\partial x - A(E)\, \partial n/\partial t \qquad (9.3-10)$$

where W, B and A are coefficients that would be determined from experimental or Monte Carlo data. Thornber claimed that Eqn. (9.3-10)

could be used to characterize the current even in the presence of sharp electric field gradients (and thus velocity overshoot). Once again though, the determination of W, B and A is not straightforward. Kizilyalli and Hess [Kiz87] have reported the determination of these coefficients for GaAs.

Thornber mentioned in that paper the he would present a detailed derivation/justification of the above equation elsewhere. But this author has not found the derivation published anywhere. Also, Blakey et al [Bla88] devised a method of testing the validity of semi-empirical transport models, and found that Thornber's equation would not be valid for electric fields below 10 kV/cm in GaAs. If this is true, then Thornber's equation will have very limited usage, since fields in the range of 1-10 kV/cm are routinely found in GaAs junction space charge layers.

CHAPTER 10
CONCLUSIONS

This doctoral work began with the goal of deriving a comprehensive analytical model for heterojunction bipolar transistors, that would include all the phenomena occurring in these transistors. But as time went by, it became increasingly clear that such a goal was too ambitious. Therefore the lesser goal of modeling some of the various phenomena individually was adopted. Thus the models discussed in chapters 5 through 8 came into being; chapter 9 described the issues that were addressed but remain unsolved. Despite the unfeasibleness of an all-comprehensive model, by the end of chapter 7 one does see the emergence of a semi-comprehensive model that is a charge-control model (thus applicable to HBTs with arbitrary doping profiles in the base, and to both single and double heterojunction transistors), that incorporates the space charge region recombination current, and that also accounts for high current density phenomena in the collector. Thus given the constraints that accompanied this work (time, facilities and the very infeasibilty of analytical modeling of the phenomena involved - velocity overshoot being a good example of this), one could say that the original goal of an all-comprehensive model was, at least partially, achieved.

There were some issues that were not addressed in this work; we mention these here for possible future work. They are listed in no particular order.

(1) Heavy doping effects: since the base of an HBT can be, and usually is, very heavily doped, these effects need to be modeled and incorporated into the transistor model. Right now, little data is available on the effects of heavy dopings in compound semiconductors [Sai89, Lun90], and thus analytical modeling has not been paid much attention. Some general references on modeling of devices having heavily doped or degenerate regions are [Marsh78a, Marsh78b, Mar79a, Vli80, Lunds81, Ben85, Marsh84, Mar87, Mar88, Cue89, Key89, Lyo89c].

(2) Energy gap grading in the base: such grading has been used to improve base transit times. Although modeling has been done [Moh90, Zha90], and computer [Maz86a, Azo89] and experimental [Haye83a, Mil83, Ito85, Mal85, Yam86b, Haye88] results also reported, the incorporation of this effect into an analytically comprehensive formulation remains to be done. We note that some of the numerical and experimental work considered the effects of hot electron transport in the base also.

(3) Tunneling-recombination currents: we saw in Sec. 3.2 that a multistep tunneling-recombination current is often a dominant current flow mechanism in heterojunctions. Little effort has been made in modeling these currents [Rib66a, Rib66b, Don66, Hal75, Sar77, Shi83], perhaps due to the complexity of the tunneling process.

(4) Surface recombination is a major source of current gain degradation in compound semiconductor HBTs as we stated at the end of chapter 6 (see Sec. 6.5 for a list of references). Models for surface recombination are few [Sah62, Henr78a, Henr78b], and in this author's opinion, far from adequate.

(5) The "alloy-barrier" effect: we mentioned this effect in Sec. 7.5 [Tiw88]. If this effect does exist, it would need analytical modeling.

(6) Optimum emitter grading: the thickness over which the emitter energy gap is graded determines the injection characteristics of the junction and also the space charge region (SCR) recombination characteristics. Optimal emitter grading will optimize the injection efficiency and the base transport factor, and minimize the SCR recombination current. Some numerical and experimental work has been done on the subject [Haye83b, Cha85c, Yos85, Tai86, Das88, Mey88, Tak88], but not analytical modeling.

(7) Collector-emitter offset voltage: In heterojunction transistors, the turn-on voltages for the emitter and collector junctions can be significantly different, since the two junctions have often disparate make-ups. The difference between the two turn-on voltages has been termed the offset voltage - ΔV_{CE}. A significant ΔV_{CE} can have repercussions on circuit applications where the transistor is required to be operated in both the forward and the reverse modes. Thus this subject has been studied in the literature [Haye84, LeeS84, Cha85b, Dan85, LeeS85, Won98], but unfortunately, this author is not very familiar with the issue.

In conclusion, the detailed literature review, interspersed with short tutorials, the many new analytical models proposed, and the myriad issues discussed in this dissertation should contribute to the existing literature on heterojunction bipolar transistor modeling.

REFERENCES

Abd79   A.Abdullaev, D.Z.Garbuzov, A.N.Ermakova and M.K.Trukan, "Estimates of the nonradiative recombination velocity at heterojunction in AlGaAs structures," Soviet Phys.-Semiconductors, 13, 1015, 1979.

Ada79   M.J.Adams and A.Nussbaum, "A proposal for a new approach to heterojunction theory," Solid State Electron., 22, 783, 1979.

Adv62   G.T.Advani, J.G.Gottling and M.S.Osman, "Thin film triode research," Proc. IRE, 50, 1530, 1962.

Ahf73   Zh.I.Ahferov, F.A.Akhmedov, V.I.Korol'kov and V.G.Nikitin, "Phototransistor utilizing a GaAs-AlAs heterojunction," Sov. Phys.-Semiconductors, 7, 780, 1973.

Ahf74   Zh.I.Ahferov, S.G.Konnikov, V.I.Korol'kov, V.B.Smirnov, D.N.Tret'yakov and A.A.Yakovenko, "Possible method for estimating the influence of the interface in heterojunctions based on solid solutions," Soviet Phys.-Semiconductors, 7, 952, 1974.

Ald82   A.W.Alden and A.R. Boothroyd, "Modeling of VHF and microwave power transistors operating in quasi-saturation," Solid State Electron., 25, 723, 1982.

And62   R.L.Anderson, "Experiments on Ge-GaAs heterojunctions," Solid State Electron., 5, 341,1962.

Ank82a   D.Ankri and L.F.Eastman, "GaAlAs-GaAs ballistic heterojunction biploar transistor," Electron. Lett., 18, 750, 1982.

Ank82b   D.Ankri, A.Scavennec, C.Besombes, C.Courbet, F.Heliot and J.Riou, "Diffused epitaxial GaAlAs-GaAs heterojunction bipolar transistor for high-frequency operation," Appl. Phys. Lett., 40, 816, 1982.

Ank83a   D.Ankri, W.J.Schaff, P.Smith and L.F.Eastman, "High-speed GaAlAs-GaAs heterojunction bipolar transistors with near ballistic operation," Electron. Lett., 19, 147, 1983.

Ank83b   D.Ankri, W.J.Schaff, C.E.C.Wood and L.F.Eastman, "GaAlAs/GaAs heterojunction bipolar transistors with abrupt emitter base

interface for ballistic operation," Gallium Arsenide and Related Compounds, p. 431, 1983.

Ant88   P. Antognetti and G. Massobrio (Ed.s), Semiconductor Device Modeling with SPICE, New York: McGraw-Hill, 1988.

Asb89a   P.M.Asbeck, "Heterojunction bipolar transistors: status and directions," Proc. Bipolar Circuits and Tech. meeting, p. 65, 1989.

Asb89b   P.M.Asbeck, M-C.F.Chang, J.A.Higgins, N.H.Sheng, G.J.Sullivan, K.C.Wang, "GaAlAs/GaAs heterojunction bipolar transistors: Issues and prospects for application," IEEE Trans. Electron Devices, ED-36, 2032, 1989.

Asbe82   P.M.Asbeck, D.L.Miller, R.Asatourian and C.G.Kirkpatrick, "Numerical simulation of GaAs/GaAlAs heterojunction bipolar transistors," IEEE Electron Device Lett., EDL-3, 403, 1982.

Azo89   E.M.Azoff, "Energy transport numerical simulation of graded AlGaAs/GaAs heterojunction bipolar transistors," IEEE Trans. Electron Devices, ED-36, 609, 1989.

Bai80   J.P.Bailbe, A.Marty, P.H.Hiep and E.G.Rey, "Design and fabrication of high-speed GaAlAs/GaAs heterojunction transistors," IEEE Trans. Electron Devices, 27, 1160, 1980.

Bai85   J.P.Bailbe, A.Marty, G.Rey, J.Tasselli and A.Bouyahyaoui, "Electrical behavior of double heterojunction NpN GaAlAs/GaAs/GaAlAs bipolar transistors," Solid State Electron., 28, 627, 1985.

Bar36   J.Bardeen, "Theory of the work function II: The surface double layer," Phys. Rev., 49, 653, 1936.

Bar47   J.Bardeen, "Surface states and rectification at a metal-semiconductor contact," Phys. Rev., 71, 717, 1947.

Bat86   J.Batey and S.L.Wright, "Energy band alignment in GaAs:(Al,Ga)As heterostructures: The dependence on alloy composition," J. Appl. Phys., 59, 200, 1986.

Bate85   J.Batey, S.L.Wright and D.J.DiMaria, "Energy band-gap discontinuities in GaAs-(Al,Ga)As heterojunctions," J. Appl. Phys., 57, 484, 1985.

Bau83   R.S.Bauer, P.Zurcher and H.W.Sang, "Inequality of semiconductor heterojunction conduction band edge discontinuity and electron affinity difference," Appl. Phys. Lett., 43, 663, 1983.

Bea68   J.R.A. Beale and J.A.G. Slatter, "The equivalent circuit of a transistor with a lightly doped collector operating in saturation," Solid State Electron., 11, p. 241, 1968.

Ben67   A.J.Bennett and C.B.Duke, "Metallic interfaces II. Influence of the exchange-correlation and lattice potentials," Phys. Rev., 162, 578, 1967.

Ben85   H.S.Bennett, "Heavy doping effects on bandgaps, effective intrinsic carrier concentrations and carrier mobilities and lifetimes," Solid State Electron., 28, 193, 1985.

Ben86   R.Benumof and J.Zoutendyk, "Theoretical values of various parameters in the Gummel-Poon model of a bipolar junction transistor," J. Appl. Phys., 59, 636, 1986.

Bet42   H.A.Bethe, "Theory of the boundary layer of crystal rectifiers," MIT Radiation Lab Report 43-12, 1942.

Bie85a   J.J.H.van den Biesen, "p-n junction capacitances, part I: the depletion capacitance," Philips J. Res., 40, 88, 1985.

Bie85b   J.J.H.van den Biesen, "p-n junction capacitances, part II: the neutral capacitance," Philips J. Res., 40, 103, 1985.

Bie86   J.J.H.van den Biesen, "A simple regional analysis of transit times in bipolar transistors," Solid State Electron., 29, 529, 1986.

Bie88   J.J.H.van den Biesen and T.Toyabe, "Comparison of methods to calculate capacitances and cutoff frequencies from DC and AC simulations on bipolar devices," IEEE Trans. CAD ICs, 7, 855, 1988.

Bla88   P.A.Blakey, S.A.Burdick and P.A.Sandborn, "On the use of Thornber's augmented drift-diffusion equation for modeling GaAs devices," IEEE Trans. Electron Devices, ED-35, 1991, 1988.

Blo70   K.Blotekjaer, "Transport equations for electrons in two valley semiconductors," IEEE Trans. Electron Devices, ED-17, 38, 1970.

Boc74   L.V.Bochkaryova and A.V.Simashkevich, "Electrical properties of ZnTe-ZnSe heterostructures," Thin Solid Films, 20, 329, 1974.

Bow73   D.L.Bowler and F.A.Lindholm, "High current regimes in transistor collector regions," IEEE Trans. Electron Devices, 20, 257, 1973.

Cap85   F.Capasso, A.Y.Cho, K.Mohammed and P.W.Foy, "Doping interface dipoles: Tunable heterojunction barrier heights and band edge disontinuities by molecular beam epitaxy," Appl. Phys. Lett., 46, 664, 1985.

Car79   H.C.Card, "Electrostatic effects of interface states on carrier transport in semiconductor heterojunctions," J. Appl. Phys., 50, 2822, 1979.

Cha85a   N.Chand, R.Fisher, J.Klem and H.Morkoc, "Current-voltage characteristics of n-AlGaAs/p-GaAs heterojunction diodes," J. Vac. Sci. Tech. B, 4, 605, 1985.

Cha85b   N.Chand, R.Fischer and H.Morkoc, "Collector-emitter offset voltage in AlGaAs/GaAs heterojunction bipolar transistors," Appl. Phys. Lett., 47, 313, 1985.

Cha85c   N.Chand and H.Morkoc, "Doping effects and compositional grading in $Al_xGa_{1-x}As$/GaAs heterojunction bipolar transistors," IEEE Trans. Electron Devices, ED-32, 1064, 1985.

Chan79   A.Chandra and L.F.Eastman, "Rectification at n-GaAs-n-GaAlAs heterojunctions grown by liquid phase epitaxy," J. Vac. Sci. Tech., 16, 1364, 1979.

Chan80   A.Chandra and L.F.Eastman, "A study of the conduction properties of a rectifying n-GaAs - n(Ga,Al)As heterojunction," Solid State Electron., 23, 599, 1980.

Chan86   C.S. Chang and H.R. Fetterman, "Electron drift velocity vs. electric field in GaAs," Solid State Electron., 29, 1295, 1986.

Chang70   C.Y.Chang and S.M.Sze, "Carrier transport across metal-semiconductor barriers," Solid State Electron., 13, 727, 1970.

ChangY67   Y.F.Chang, "The capacitance of p-n junctions," Solid State Electronics, 10, 281, 1967.

Che88   M.K.Chen, F.A.Lindholm and B.S.Wu, "Comparison and extension of recent one-dimensional bipolar transistor models," IEEE Trans. Electron Devices, ED-35, 1096, 1988.

Cho75   J.Choma, "A process-oriented model for the simulation of base pushout in integrated bipolar devices," IEEE Trans. Electron Devices, ED-22, 1079, 1975.

Chr66   S.G.Christov, "General theory of electron emission from metals," Physica Stat. Solidi, 17, 11, 1966.

Chr67   S.G.Christov, "Unified theory of thermionic and field
           emission from semiconductors," Physica Stat. Solidi, 21, 159,
           1967.

Chy61   A.G.Chynoweth, W.L.Feldman and R.A.Logan, "Excess tunnel
           current in silicon Esaki junctions," Phys. Rev., 121, 684,
           1961.

Cla69   L.E. Clark, "Characteristics of two-region saturation
           phenomena," IEEE Trans. Electron Devices, ED-16, 113, 1969.

Con66   J.W.Conley, C.B.Duke, G.D.Mahan and J.J.Tiemann, "Electron
           tunneling in metal-semiconductor barriers," Phys. Rev., 150,
           466, 1966.

Con85   E.Constant, "Non-steady-state transport in semiconductors in
           perspective with submicrometer devices," in book:
           L.Reggiani(ed.), Hot-electron Transport in Semiconductors,
           Springer-Verlag, Berlin, 1985.

Coo82   R.K.Cook and J.Frey, "Teo-dimensional numerical sumulation of
           energy transport in Si and GaAs MESFETs," IEEE Trans.
           Electron Devices, ED-29, 970, 1982.

Cro88   C.R.Crowell and M.Hafizi, "Current transport over parabolic
           potential barriers in semiconductor devices," IEEE Trans.
           Electron Devices, ED-35, 1087, 1988.

Crow69   C.R.Crowell and V.L.Rideout, "Normalized thermionic-field (T-
           F) emission in metal-semiconductor (Schottky) barriers,"
           Solid State Electron., 12, 89, 1969.

Crowe66   C.R.Crowell and S.M.Sze, "Current transport in metal-
           semiconductor barriers," Solid State Electron., 9, 1035,
           1966.

Cue89   A.Cuevas and M.A.Balbuena, "Review of analytical models for
           the study of highly doped regions of silicon devices," IEEE
           Trans. Electron Devices, ED-36, 553, 1989.

Dan85   J.Dangla, E.Caquot, C.Dubon, M.Campana, R.Azoulay,
           F.Alexandre, J.Leivin and J.F.Palmier, "Modeling of DC
           characteristics of heterojunction bipolar transistor
           processed with MBE or MOCVD techniques," Physica B & C,
           129B+C, 366, 1985.

Das88   A.Das and M.S.Lundstrom, "Numerical study of emitter-base
           junction design for AlGaAs/GaAs heterojunction bipolar
           transistors," IEEE Trans. Electron Devices, ED-35, 863, 1988.

Daw75   A.N.Daw, T.Sinha, A.K.Dutta and N.K.D.Choudhury, "The effect of spatial variation of mobility on the input conductance and base charging capacitance of a diffused base transistor," Int. J. Electron., 38, 793, 1975.

Don66   J.P.Donnelly and A.G.Milnes, "Current/voltage characteristics of p-n Ge-Si and Ge-GaAs heterojunctions," Proc. IEE, 113, 1468, 1966.

Dum72   W.P.Dumke, J.M.Woodall and V.L.Rideout, "GaAs-GaAlAs heterojunction transistor for high frequency operation," Solid State Electron., 15, 1339, 1972.

Dun73   W.Duncan, J.Lamb, K.G.McIntosh and A.R.Smellie, "Current-voltage characteristics of p-Ge/n-CdS heterojunction diodes," Appl. Phys. Lett., 23, 330, 1973.

Ebe54   J.J.Ebers and J.L.Moll, "Large signal behavior of junction transistors," Proc. IRE, 42, 1761, 1954.

Ede82   R.C.Eden, "Comparison of GaAs device approaches for ultrahigh-speed VLSI," Proc. IEEE, 70, 5, 1982.

End86   P.Enders, "Interface recombination velocity for a barrier at a p-n heterojunction," Phys. Stat. Solidi A, 94, K85, 1986.

Eng86   W.L.Engl (ed.), Process and Device Modeling, Springer-Verlag, Amsterdam, 1986.

Enq87   P.M.Enquist, L.P.Ramberg and L.F.Eastman, "Comparison of compositionally graded to abrupt emitter-base junctions used in the heterojunction bipolar transistor," J. Appl. Phys., 61, 2663, 1987.

Ett76   M.Ettenberg and H.Kressel, "Interfacial recombination at (AlGa)As/ GaAs heterojunction structure," J. Appl. Phys., 47, 1536, 1976.

Ett77   M.Ettenburg, C.J.Nuese and G.H.Olsen, "Interfacial recombination velocity determination in InGaP/GaAs," J. Appl. Phys., 48, 1288, 1977.

Eva87   S.Evans, J.Delaney, C.Fuller, D.Boone, C.Dubberley, J.Hoff, J.Stidham, B.de la Torre, M.Vernon and M.Wdosik, "GaAs HBT LSI/VLSI fabrication technology", GaAs IC Symposium Technical Digest, 1987, p.109.

Flo77   F.Flores and C.Tejedor, "A simple approach to covalent surgaces," Journal de Physique, 38, 949, 1977.

For86   A.Fortini and S.Munoglu, "High level boundary conditions in the space charge region of a PN junction," Solid State Electron., $\underline{29}$, 1107, 1986.

Fos81   J.G.Fossum and M.A.Shibib, "An analytic model for minority carrier transport in heavily-doped regions of silicon devices," IEEE Trans. Electron Devices, $\underline{ED-28}$, 1018, 1981.

Fos86   J.G.Fossum and S.Veeraraghavan, "Partitioned-charge-based modeling of bipolar transistors for non-quasi-static circuit simulation," IEEE Electron Device Lett., $\underline{7}$, 652, 1986.

Fow28   R.H.Fowler and L.Nordheim, "Electron emission in intense electric fields," Proc. Roy. Soc.(London), $\underline{A119}$, 173, 1928.

Fre76   W.R.Frensley and H.K.Kroemer, "Prediction of semiconductor heterojunction discontinuities from bulk band structures," J.Vac.Sci.Tech., $\underline{13}$, 810, 1976.

Fre77   W.R.Frensley and H.Kroemer, "Theory of the energy band lineup at an abrupt semiconductor heterojunction," Phys. Rev. B, $\underline{16}$, 2642, 1977.

Gar79a   C. M. Garner, C. Y. Su, Y. D. Shen, C. S. Lee, G. L. Pearson and W. E. Spicer, "Interface studies of $Al_xGa_{1-x}As$-GaAs heterojunctions," J. Appl. Phys., $\underline{50}$, 3383, 1979.

Gar79b   C. M. Garner, C. Y. Su and W. E. Spicer, "Interface studies of AlGaAs-GaAs heterojunctions," J. Vac. Sci. & Tech., $\underline{16}$, 1364, 1979.

Gas73   P.A.Gashin and A.V.Simashkevich, "ZnTe-CdSe heterojunctions I:Electrical properties," Phys. Stat. Sol. A, $\underline{19}$, 379, 1973.

Gau75   S.P.Gaur, "Quasisaturation region operation of $n-p-n^--n$ power transistors," Electron. Lett., $\underline{11}$, 446, 1975.

Get78   I. Getreu, Modeling the Bipolar Transistor, Elsevier Scientific Pub. Co., New York, 1978.

Gia74   L.J.Giacolletto, "A model for d.c. characterization of bipolar transistors," Int. J. Electron., $\underline{36}$, 719, 1974.

Gor80   P.P.Gorbik, V.N.Komashchenko and G.N.Fedorus, "Mechanisms of current flow and energy band diagram of copper sulfide-zinc sulfide heterojuctions," Sov. Phys.-Semiconductors, $\underline{14}$, 753, 1980.

Gra73   H.C.de Graaff, "Collector models for bipolar transistors," Solid State Electron., $\underline{16}$, 587, 1973.

Gra90   H.C.de Graaff and F.M.Klaassen, <u>Compact Transistor Modeling for Circuit Design</u>, Springer-Verlag, NEw York, 1990.

Graa85   H.C.de Graaff and W.J.Kloosterman, "New formulation of the current and charge relations in bipolar transistor modeling for CACD purposes," IEEE Trans. Electron Devices, <u>ED-32</u>, 2415, 1985.

Gri79   Z.S.Gribnikov and Yu.S.Mel'nikova, "Two types of behavior of non-equilibrium carriers in p-n heterojunctions at high current densities," Sov. Phys.-Semiconductors, <u>13</u>, 364, 1979.

Gri84   A.A.Grinberg, M.S.Shur, R.J.Fischer and H.Morkoc, "An investigation of the effect of graded layers and tunneling on the performance of AlGaAs/GaAs heterojunction bipolar transistors," IEEE Trans. Electron Devices, <u>31</u>, 1758, 1984.

Gru80   H.L.Grubin, "Hot electron effects in semiconductor devices," in book: D.K.Ferry, R.J.Barker and C.Jacoboni (ed.s), <u>Physics of Nonlinear Transport in Semiconductors</u>, Plenum, New York, 1980.

Gum70a   H. K. Gummel, "A charge control relation for bipolar transistors," Bell Syst. Tech. I., <u>49</u>, 115, 1970.

Gum70b   H.K.Gummel and H.C.Poon, "An integral charge control model for the bipolar transistor," Bell Syst. Tech. J., <u>49</u>, 827, 1970.

Gumm67   H.K.Gummel and D.L.Scharfetter, "Depletion layer capacitance of p^+n step junctions," J. Appl. Phys., <u>38</u>, 2148, 1967.

Hal75   B.Halil and K.C.Kao, "Tunneling and recombination processes in heterojunctions with interface states," Int.J.Electron., <u>38</u>, 1, 1975.

Hall52   R.N.Hall, "Electron-hole recombination in Germanium," Phys. Rev., <u>87</u>, 387, 1952.

Hall87   S.Hall and W.Eccleston, "The heterojunction bipolar transistor: an estimate of its potential for digital applications," J. Inst. Electron. Radio Eng., <u>57</u>, 29, 1987.

HallW73 W.F.Hall, "Derivation of current-voltage characteristics for graded heterojunctions," Electron. Lett., <u>9</u>, 548, 1973.

Har77   W.A.Harrison, "Elementary theory of heterojunctions," J. Vac. Sci. Tech., <u>14</u>, 1016, 1977.

Hau71   J.R.Hauser, "Boundary conditions at p-n junctions," Solid State Electron., <u>14</u>, 133, 1971.

Hay87   N.Hayama, A.Okamoto, M.Madihian and K.Honjo, "Subμm fully
             self-aligned AlGaAs/GaAs heterojunction bipolar transistors,"
             IEEE Electron Device Lett., EDL-8, 246, 1987.

Haye83a   J.R.Hayes, F.Capasso, A.C.Gossard, R.J.Malik and W.Wiegmann,
             "Bipolar transistor with graded band-gap base," Electron.
             Lett., 19, 410, 1983.

Haye83b   J.R.Hayes, F.Capasso, R.J.Malik, A.C.Gossard and W.Wiegmann,
             "Optimum emitter grading for heterojunction bipolar
             transistors," Appl. Phys. Lett., 43, 949, 1983.

Haye84   J.R.Hayes, A.C.Gossard and W.Wiegmann, "Collector/emitter
             offset voltage in double heterojunction bipolar transistors,"
             Electron. Lett., 20, 766, 1984.

Haye88   J.R.Hayes and J.P.Harbison, "Hot electron transport in a
             graded bandgap heterojunction bipolar transistor," Appl.
             Phys. Lett., 53, 490, 1988.

Hayes86   J.R.Hayes, A.F.J.Levi, A.C.Gossard and J.H.English, "Base
             transport dynamics in a heterojunction bipolar transistor,"
             Appl. Phys. Lett., 49, 1481, 1986.

Hea79   E.L.Heasell, "Boundary conditions at p-n junctions," Solid
             State Electron., 22, 853, 1979.

Heb88   F.Hebert and D.J.Roulston, "Voltage and current dependent
             model for the base resistance of bipolar transistors," IEEE
             Trans. Electron Devices, ED-35, 1696, 1988.

Hei87   H.Heinrich, "Potential steps in semiconductor
             heterostructures: Their importance, determination and
             prediction," Acta Phys. Pol. A, A71, 125, 1987.

Hei88   H.Heinrich, "Band-edge offsets in semiconductor
             heterojunctions," in book: G.Ferenczi and F.Beleznay(ed.s),
             New Developments in Semiconductor Physics. Springer-Verlag,
             Berlin, 1988.

Hen84   H.K.Henisch: Semiconductor Contacts: An Approach to Ideas and
             Models. Clarendon Press, Oxford, 1984.

Henr77a   C.H.Henry and R.A.Logan, "The origin of large dark spots in
             AlGaAs-GaAs heterostructure photoluminescence," Appl. Phys.
             Lett., 31, 203, 1977.

Henr77b   C.H.Henry and R.A.Logan, "Nonradiative 'large dark spots' in
             Al$_x$Ga$_{1-x}$As-GaAs heterostructures," IEEE Trans. Electron
             Devices, ED-24, 1215, 1977.

Henr78a   C.H.Henry, R.A.Logan and F.R.Merritt, "The effect of surface
           recombination on current in $Al^xGa^{1-x}As$ heterojunctions," J.
           Appl. Phys., 49, 3530, 1978.

Henr78b   C.H.Henry and R.A.Logan, "Effect of interface recombination
           at $Al^xGa^{1-x}As$ p-n junction perimeters on photoluminescence
           and current," J. Vacuum Science and Tech., 15, 1471, 1978.

Hig88   J.A. Higgins, "Heterojunction bipolar transistors for high
           efficiency power amplifiers," GaAs IC Symposium Digest, p.
           33, 1988.

Hil86   A.J.Hill and P.H.Ladbrooke, "Dependence of conduction band
           discontinuity on Al mole fraction in GaAs/AlGaAs
           heterojunctions," Electron. Lett., 22, 218, 1986.

Hor89   K.Horio, Y.Iwatsu and H.Yanai, "Numerical simulation of
           AlGaAs/GaAs heterojunction bipolar transistors with varioius
           collector parameters," IEEE Trans. Electron Devices, 36, 617,
           1989.

Hor90   K.Horio and H.Yanai, "Numerical modeling of heterojunctions
           including the thermionic emission mechanism at the
           heterojunction interface," IEEE Trans. Electron Devices, ED-
           37, 1093, 1990.

Hov69   H.J.Hovel and A.G.Milnes, "ZnSe-Ge heterojunction
           transistors," IEEE Trans. Electron Devices, 16, 766, 1969.

How73   D.S.Howarth and D.L.Feucht, "Barrier determinations on Ge-
           AlGaAs and GaAs-AlGaAs p-n heterojunctions," Appl. Phys.
           Lett., 23, 365, 1973.

How78   P.L.Hower and W.G.Einthoven, "Emitter current crowding in
           high voltage transistors," IEEE Trans. Electron Devices, ED-
           25, 465, 1978.

Hu89   J.Hu, K.Tomizawa and D.Pavlidis, "Monte Carlo approach to
           transient analysis of HBTs with different collector designs,"
           IEEE Electron Device Lett., 10, 55, 1989.

Hur88   G.A.M.Hurkx, "A new approach to AC characterization of
           bipolar transistors," Solid State Electron., 31, 1269, 1988.

Ihm79   J.Ihm and M.L.Cohen, "Self-consistent calculation of the
           electronic structure of the (110) GaAs-ZnSe interface," Phys.
           Rev. B, 20, 729, 1979.

Ima87   T.Imai, S.Fuke, A.Takeuchi, T.Uchida and K.Kuwahara,
           "Electrical behavior of Np AlGaAs/GaAs heterojunctions with

high acceptor concentrations," Solid State Electron., <u>30</u>, 865, 1987.

Ito85   H.Ito, T.Ishibashi and T.Sugeta, "Current gain enhancement in graded base AlGaAs/GaAs HBTs associated with electron drift motion," Japanese J. Appl. Phys. 2, <u>24</u>, L241, 1985.

Jad69   D.K.Jadus and D.L.Feucht, "The realization of a GaAs-Ge wide band gap emitter transistor," IEEE Trans. Electron Devices, <u>16</u>, 102, 1969.

Jen58   D.A.Jenny, "The status of transistor research in compound semiconductors," Proc. IRE, <u>46</u>, 959, 1958.

Jeo87   J.Jeong, T.Schlesinger and A.G.Milnes, "Consideration of discrete interface traps in InGaAs/GaAs heterojunctions," IEEE Trans. Electron Devices, <u>ED-34</u>, 1911, 1987.

Jeon87   H. Jeong and J.G. Fossum, "Physical modeling of high current transients for bipolar transistor circuit simulation," IEEE Trans. Electron Devices, <u>ED-34</u>, 898, 1987.

Jeo89   H. Jeong and J.G. Fossum, "A charge-based large-signal bipolar transistor model for device and circuit simulation," IEEE Trans. Electron Devices, <u>ED-36</u>, 124, 1989.

Jer76   J.Jerhot and V.Snejdar, "The influence of interface states on an isotype heterojunction between two wide-gap semiconductors," Phys. Stat. Solidi A, <u>34</u>, 505, 1976.

Jos88   H.F.F.Jos, "Bipolar transistor: two-dimensional effects on current gain and base transit time," Solid State Electron., <u>31</u>, 1715, 1988.

Kan80   B.D.Kandilarov, V.Detcheva and M.T.Primatarowa, "Energy band profile and interface potential step in the theory of heterojunctions," Surf. Sci., <u>99</u>, 174, 1980.

Kand79   B.D.Kandilarov and M.T.Primatarowa, "Energy band profile and interface states in semiconductor heterojunctions," J. Phys. C, <u>12</u>, L463, 1979.

Kandi73 B.D.Kandilarov, M.G.Tashkova and P.C.Petrova, "On the Aerts model in the theory of interface states," C.R. Acad. Bulg. Sci., <u>26</u>, 159, 1973.

Kat85   A.D.Katnani, P.Chiaradia, H.W.Sang,Jr. and R.S.Bauer, "Irrelevance of interface defects to heterojunction band offsets," J. Electron. Mater., <u>14</u>, 25, 1985.

Kat87 R.Katoh, M.Kurata and J.Yoshida, "A self-consistent particle simulation for (AlGa)As/GaAs HBTs," IEDM Tech. Digest, 1987, p.248.

Kat89a R.Katoh, M.Kurata, "Self-consistent partical simulation for AlGaAs/GaAs heterojunction bipolar transistors under high bias conditions," IEEE Trans. Electron Devices, ED-36, 2122, 1989.

Kat89b R.Katoh, M.Kurata and J.Yoshida, "Self-consistent particle simulation for AlGaAs/GaAs HBTs with improved base-collector structures," IEEE Trans. Electron Devices, ED-36, 846, 1989.

Kau78 D.K.Kaushik and S.K.Chattopadhyaya, "An analysis of base transit time of diffused base transistors," Indian J. Pure and Appl. Phys., 16, 770, 1978.

Ken38 E.H.Kennard, Kinetic Theory of Gases, McGraw-Hill, New York, 1938.

Key89 R.W.Keyes, "Potential and junctions in degenerate semiconductors," Solid State Electron., 32, 159, 1989.

Kim85 B.Kim, H.Q.Tseng, S.K.Tiku and H.D.Shih, "AlGaAs/GaAs heterojunction bipolar power transistors," Electron. Lett., 21, 258, 1985.

Kir62 C.T.Kirk, "A theory of transistor cutoff frequency (f_T) falloff at high current densities", IRE Trans. Electron Devices, ED-9, 164, 1962.

Kiz87 I.C.Kizilyalli and K.Hess, "Simplified device equations and transport coefficients for GaAS device modeling," IEEE Trans. Electron Devices, 34, 2352, 1987.

Klo87 H.Klose and A.W.Wieder, "The transient integral charge control relation - a novel formulation of the current in a bipolar transistor," IEEE Trans. Electron Devices, 34, 1090, 1987.

Kon75 M.Konogai and K.Takahashi, "GaAlAs-GaAs heterojunction transistors with high injection efficiency," J. Appl. Phys., 46, 2120, 1975.

Kra78 S.Kratzer and J.Fray, "Transient velocity characteristics of electrons in GaAs with Γ-L-X conduction band ordering," J. Appl. Phys., 49, 4064, 1978.

Kro59 H.Kroemer, "Theory of a wide-gap emitter transistor," Proc. IRE, 45, 1535, 1959.

Kro75   H.Kroemer, "Problems in the theory of heterojunction discontinuities," Crit. Rev. Solid State Sci., 5, 555, 1975.

Kro83a   H.Kroemer, "Theory of heterojunctions: A critical review," in L.L.Chang (ed.): Proc. NATO Advanced Study Institute on Molecular Beam Epitaxy and Heterostructures, Erice, Sicily, 1983. Martinus Nijhoff, Netherlands.

Kro83b   H.Kroemer, "Critique of two recent theories of heterojunction lineups," IEEE Electron Device Lett., EDL-4, 25, 1983.

Kro84   H.Kroemer, "Barrier control and measurements: Abrupt semiconductor heterojunctions," J. Vac. Sci. Tech. B, 2, 433, 1984.

Kro85a   H.Kroemer, "Determination of heterojunction band offsets by capacitance-voltage profiling through nonabrupt isotype heterojunctions," Appl. Phys. Lett., 46, 504, 1985.

Kro85b   H.Kroemer, "Two integral relations pertaining to the electron transport through a bipolar transistor with a nonuniform energy gap in the base region," Solid State Electron., 28, 1101, 1985.

Kroe78   H.Kroemer, W.Y.Chien, H.C.Casey and A.Y.Cho, "Photocollection efficiency and interface charges of MBE-grown abrupt p(GaAs)-N(AlGaAs) heterojunctions," Appl. Phys. Lett., 33, 749, 1978.

Kul85   G.M. Kull, L.W. Nagel, S.-W. Lee, P. Lloyd. E.J. Prendergast, and H. Dirks, "A unified circuit model for bipolar transistors including quasi-saturation effects," IEEE Trans. Electron Devices, ED-32, 1103, 1985.

Kum75   R. Kumar and L.P. Hunter, "Prediction of f_t and h_e at high collector currents," IEEE Trans. Electron Devices, ED-22, 1031, 1975.

Kur84   M.Kurata and J.Yoshida, "Modeling and characterization for high-speed GaAlAs-GaAs n-p-n heterojunction bipolar transistors," IEEE Trans. Electron Devices, ED-31, 467, 1984.

Kwe76   A.H.Kwesah and M.K.McPhun, "New π equivalent circuit for high-frequency bipolar transistors," Electron. Lett., 12, 334, 1976.

Lan77   D.V.Lang and R.A.Logan, "A search for interface states in an LPE GaAs/Al_xGa_{1-x}As heterojunction," Appl. Phys. Lett., 31, 683, 1977.

Lee85   R.J.Lee, "Some comments on Nussbaum's heterojunction lineup theory," IEEE Electron Device Lett., EDL-6, 130, 1985.

LeeS81   S.C.Lee and G.L.Pearson, "Rectification in $Al_xGa_{1-x}As$-GaAs N-n heterojunction devices," Solid State Electron., $\underline{24}$, 563, 1981.

LeeS84   S.C.Lee, J-N.Kau and H-H.Lin, "Origin of high offset voltage in an AlGaAs/GaAs heterojunction bipolar transistor," Appl. Phys. Lett., $\underline{45}$, 1114, 1984.

LeeS85   S.C.Lee, J-N.Kau and H-H.Lin, "Current transport across the emitter-base potential spike in AlGaAs/GaAs heterojunction bipolar transistors," J. Appl. Phys., $\underline{58}$, 890, 1985.

LeeS86   S.C.Lee and H.-H.Lin, "Transport theory of the double heterojunction bipolar transistor based on current balancing concept," J. Appl. Phys., $\underline{59}$, 1688, 1986.

LeeW89   W-S.Lee, D.Ueda, T.Ma, Y-C.Pao and J.S.Harris, "Effect of emitter-base spacing on the current gain of AlGaAs/GaAs heterojunction bipolar transistors," IEEE Electron Device Lett., $\underline{EDL-10}$, 200, 1989.

Lin85   H.-H.Lin and S.-C.Lee, "Direct measurement of the potential spike energy in AlGaAs/GaAs single heterojunction bipolar transistors," IEEE Electron Device Lett., $\underline{6}$, 431, 1985.

Lin88   F.A.Lindholm and J.J.Liou, "Quasi-static capacitance of p/n junction space-charge layers by Leibnitz rule," J. Appl. Phys., $\underline{63}$, 561, 1988.

Lio89   J.J. Liou, "Investigation of high-current effects on the current gain of $Al_xGa_{1-x}As$/GaAs abrupt heterojunction bipolar transistors," Solid State Electron., $\underline{32}$, 169, 1989.

Liou88   J.J.Liou and F.A.Lindholm, "Capacitance of semiconductor p-n junction space charge layers: an overview," Proc. IEEE, $\underline{76}$, 1406, 1988.

Lio90   J.J.Liou, F.A.Lindholm and B.S.Wu, "Modeling the cutoff frequency of single-heterojunction bipolar transistors subjected to high collector-layer current," J. Appl. Phys., $\underline{67}$, 7125, 1990.

Log77   R.A.Logan and F.R.Merritt, "Origin of $n \simeq 2$ injection current in Al_xGa_{1-x}-GaAs heterojunctions," Appl. Phys. Lett., $\underline{31}$, 454, 1977.

Lun84   M. S. Lundstrom, "Boundary conditions for pn hetero-junctions," Solid State Electron., $\underline{27}$, 491, 1984.

Lun86   M.S.Lundstrom, "An Ebers-Moll model for the heterostructure transistor," Solid State Electron., $\underline{29}$, 1173, 1986.

Lun90   M.S.Lundstrom, M.E.Klausmier-Brown and M.R.Melloch, "Device-related material properties of heavily doped gallium arsenide," Solid State Electron., $\underline{33}$, 693, 1990.

Lund82   M.S.Lundstrom and R.J.Schuelke, "Modeling semiconductor heterojunctions in equilibrium," Solid State Electron., $\underline{25}$, 683, 1982. [168]

Lunds81   M.S.Lundstrom, R.J.Schwartz and J.L.Gray, "Transport equations for the analysis of heavily doped semiconductor devices," Solid State Electron., $\underline{24}$, 195, 1981.

Lup77   V.M.Lupin and P.E.Ramazanov, "Electrical properties of CdS-GaAs heterojunctions," Sov. Phys. J., 1977.

Lyo89a   T.J.deLyon, H.C.Casey, P.M.Enquist, J.A.Hutchby, and A.J.SpringThorpe, "Surface recombination current and emitter compositional grading in npn and pnp heterojunction bipolar transistors," Appl. Phys. Lett., $\underline{54}$, 641, 1989.

Lyo89b   T.J.deLyon, H.C.Casey, P.M.Enquist, J.A.Hutchby, and A.J.SpringThorpe, "Observation of radiative and nonradiative tunneling in $GaAs/Al_xGa_{1-x}As$ heterojunction bipolar transistors with compositionally graded base-emitter heterojunctions," J. Appl. Phys., $\underline{65}$, 3282, 1989.

Lyo89c   T.J.deLyon, H.C.Casey and A.J.SpringThorpe, "Nonconventional electron diffusion current in $GaAs/Al_xGa_{1-x}As$ N-p-n heterojunction bipolar transistors with heavily doped base layers," J. Appl. Phys., $\underline{65}$, 2530, 1989.

Ma89   T.Ma, W-S.Lee, J.W.Adkisson and J.S.Harris, "Effect of bulk recombination current on the current gain of GaAs/AlGaAs heterojunction bipolar transistors in GaAs-on-Si," IEEE Electron Device Lett., $\underline{10}$, 458, 1989.

Mae86   H.Maeda, "Electron transport across a semiconductor heterojunction," Japanese J. Appl. Phys. 1, $\underline{25}$, 1221, 1986.

Mal85   R.J.Malik, F.Capasso, R.A.Stall, R.A.Kiehl, R.W.Ryan, R.Wunder and C.G.Bethea, "High gain, high frequency AlGaAs/GaAs graded band-gap base bipolar transistors with a Be diffusion setback layer in the base," Appl. Phys. Lett., $\underline{46}$, 600, 1985.

Malo77   T.J.Maloney and J.Frey, "Transient and steady-state electron transport properties of GaAs and InP," J. Appl. Phys., $\underline{48}$, 781, 1977.

Mar83a   G.Margaritondo, "Microscopic investigations of semiconductor interfaces," Solid State Electron., $\underline{26}$, 499, 1983.

Mar83b G.Margaritondo, "Heterojunction parameters from a microscopic point of view," Surface Sci., <u>132</u>, 469, 1983.

Mar84 G.Margaritondo, C.Capasso, F.Patella, P.Perfetti, C.Quaresima, A.Savoia, F.Sette, "Order, disorder, and band discontinuities at ZnSe-Ge heterojunctions," J. Vac. Sci. Tech. A, <u>2</u>, 508, 1984.

Mar87 A.H.Marshak, "Transport equations for highly doped devices and heterostructures," Solid State Electron., <u>30</u>, 1089, 1987.

Mar88 A.H.Marshak, "On the nonequilibrium carrier density equations for highly doped devices and heterostructures," Solid State Electron., <u>31</u>, 1551, 1988.

Mars79 A.H.Marshak and R.Srivastava, "Law of the junction for degenerate materials with position dependent band-gap and electron affinity," Solid State Electron., <u>22</u>, 567, 1979.

Marsh78a A.H.Marshak and K.M.van Vliet, "Electrical currents in solids with position dependent band structure," Solid State Electron., <u>21</u>, 417, 1978.

Marsh78b A.H.Marshak and K.M.van Vliet, "Carrier densities and emitter efficiency in degenerate materials with position dependent band structures," Solid State Electron., <u>21</u>, 429, 1978.

Marsh80 A.H.Marshak and K.M.van Vliet, "On the separation of quasi-Fermi levels and the boundary conditions for junction devices," Solid State Electron., <u>23</u>, 1223, 1980.

Marsh84 A.H.Marshak and C.M.van Vliet, "Electrical current and carrier density in degenerate materials with nonuniform band structure," Proc. IEEE, <u>72</u>, 148, 1984.

Mart76 S.Martinuzzi, O.Mallem and T.Cabot, "Dark currents in Cu_2S-CdS single crystal heterojunctions," Phys. Stat. Solidi A, <u>36</u>, 227, 1976.

Mart88 A.Marty, J.Jamai, J.P.Vannel, N.Fabre, J.P.Bailbe, N.DUhamel, C.Dubon-Chevallier and J.Tasselli, "Fabrication and d.c. characterization of GaAlAs/GaAs heterojunction bipolar transistors," Solid State Electron., <u>31</u>, 1375, 1988.

Marty79 A.Marty, G.Rey and J.P.Bailbe, "Electrical behavior of an Npn GaAlAs/GaAs heterojunction transistor," Solid State Electron., <u>22</u>, 549, 1979.

Mat83 T.Matsumoto, P.K.Bhattacharya and M.J.Ludowise, "Interface states in GaAs/$Al_xGa_{1-x}As$ heterostructures grown by

organometallic vapor phase epitaxy," Appl. Phys. Lett., $\underline{42}$, 52, 1983.

Maz86a   C.M.Maziar, M.E.Klausmier-Brown, S.Bandyopadhyay, M.S.Lundstrom and S.Datta, "Monte Carlo evaluation of electron transport in heterojunction bipolar transistors," IEEE Trans. Electron Devices, $\underline{ED-33}$, 881, 1986.

Maz86b   C.M.Maziar, M.E.Klausmier-Brown, M.S.Lundstrom, "A proposed structure for collector transit-time reduction in AlGaAs/GaAs bipolar transistors," IEEE Electron Device Lett., $\underline{EDL-7}$, 483, 1986.

McA82   S.R.McAfee, D.V.Lang and W.T.Tsang, "Observation of deep levels associated with the $GaAs/Al_xGa_{1-x}As$ interface grown by molecular beam epitaxy," Appl. Phys. Lett., $\underline{40}$, 520, 1982.

McD88   R.J.McDonald, "Generalized partitioned-charge-based bipolar transistor modeling methodology," Electron. Lett., $\underline{24}$, 1302, 1988.

Mey87   R.G.Meyer and R.S.Muller, "Charge-control analysis of the collector-base space charge region contribution to bipolar transistor time constant τ_T," IEEE Trans. Electron Devices, $\underline{ED-34}$, 450, 1987.

Mey88   M. Meyyappan, G. Andrews and H. L. Grubin, "Numerical study of an AlGaAs/GaAs heterostructure bipolar transistor: emitter design and compositional grading," Solid State Electron., $\underline{31}$, 1611, 1988.

Mey89   M.Meyappan, G.Andrews, H.L.Grubin and J.P.Krevskovsky, "Analysis of a self-aligned AlGaAs/GaAs heterostructure bipolar transistor: steady-state and transient simulations," J. Appl. Phys., $\underline{66}$, 3348, 1989.

Mil83   D.L.Miller, P.M.Asbeck, R.J.Anderson and F.H.Eisen, "(GaAl)As/GaAs heterojunction bipolar transistors with graded composition in the base," Electron. Lett., $\underline{19}$, 367, 1983.

Mil86   A.G.Milnes, "Semiconductor heterojunction topics: introduction and overview," Solid State Electron., $\underline{29}$, 99, 1986.

Mil87   A.G.Milnes, "Heterojunctions: some knowns and unknowns," Solid State Electron., $\underline{30}$, 1099, 1987.

Miln72   A.G.Milnes and D.L.Feucht, Heterojunctions and Metal-Semiconductor Junctions, Academic Press, New York, 1972.

Moh87   S.N.Mohammad, "Emitter efficiency, transit times and current gain of bipolar transistors," Solid State Electron., <u>30</u>, 685, 1987.

Moh90   S.N.Mohammad, "Influence of doping dependent bandgap grading on electrical performance and design criteria of npn AlGaAs/Gaas abrupt heterojunction bipolar transistors," Solid State Electron., <u>33</u>, 339, 1990.

Mor88   K.Morizuka, R.Katoh, K.Tsuda, M.Asaka, N.Iizuka and M.Obara, "Electron space-charge effects on high frequency performance of AlGaAs/GaAs HBTs under high-current-density operation," IEEE Electron Device Lett., <u>9</u>, 570, 1988.

Mur56   E.L.Murphy and R.H.Good, "Thermionic emission, field emission and the transition region," Phys. Rev., <u>102</u>, 1464, 1956.

Nag89   K.Nagata, O.Nakajima, T.Nittono, Y.Yamauchi and T.Ishibashi, "Extremely small emitter (1x1 μm^2) AlGaAs/GaAs HBT's utilizing bridged base electrode structure," IEDM Tech. Dig., p. 385, Dec. 1989.

Nak85a   O.Nakajima, K.Nagata, H.Ito, T.Ishibashi and T.Sugeta, "Emiiter-base junction size effect on current gain H_{fe} of AlGaAs/GaAs heterojunction bipolar transistors," Japanese J. Appl.Phys. part 2, <u>24</u>, L596, 1985.

Nak85b   O.Nakajima, K.Nagata, H.Ito, T.Ishibashi and T.Sugeta, "Suppression of emitter size effect on current gain in AlGaAs/GaAs HBTs," Japanese J. Appl. Phys. 1, <u>24</u>, 1368, 1985.

Nan88   M.Nanba, T.Shiba, T.Nakamura and T.Toyabe, "An analytical and experimental investigation of the cutoff frequency f_T of high-speed bipolar transistors," IEEE Trans. Electron Devices, <u>ED-35</u>, 1021, 1988.

Neg88   K.J.Negus and D.J.Roulston, "Simplified modeling of delays in the emitter-base junction," Solid State Electron., <u>31</u>, 1464, 1988.

Nel78a   R.J.Nelson, "Interfacial recombination in GaAlAs/GaAs heterostructures," J. Vac. Sci. and Tech., <u>15</u>, 1475, 1978.

Nel78b   R.J.Nelson and R.G.Sobers, "Interfacial recombination velocity in GaAlAs/GaAs heterostructures," Appl. Phys. Lett., <u>32</u>, 761, 1978.

Nor28   L.Nordheim, "The effect of the image force on the emission and reflexion of electrons by metals," Proc. Roy. Soc.(London), <u>A121</u>, 626, 1928.

Not89   R.N.Nottenburg, Y.K.Chen, M.B.Panish, D.A.Humphrey and
         R.A.Hamm, "Hot electron InGaAs/InP heterostructure bipolar
         transistors with f_T of 110 GHz," IEEE Electron Device Lett.,
         EDL-10, 30, 1989.

Nott88   R.N.Nottenburg, C.J.Sandroff, D.A.Humphrey, T.H.Hollenbeck
         and R.Bhat, "Near-ideal transport in an AlGaAs/GaAs
         heterostructure bipolar transistor by $Na_2S \cdot 9H_2O$ regrowth,"
         Appl. Phys. Lett., 52, 218, 1988.

Nus84   A.Nussbaum, "Heterojunction discontinuities: the current
         position," IEEE Electron Device Lett., EDL-5, 499, 1984.

Nus85   A.Nussbaum, "Response to a set of comments on heterojunction
         discontinuities," IEEE Electron Device Lett., EDL-6, 273,
         1985.

Nuy71   W.Nuyts and R.van Overstraeten, "Analytica expressions for
         electric field and for capacitance of symmetrical and
         asymmetrical abrupt p-n junctions," J. Appl. Phys., 42, 5109,
         1971.

Oku85   H.Okumura, S.Misawa, S.Yoshida and S.Gonda, "Determination of
         the conduction band discontinuity of GaAs/AlGaAs interfaces
         by capacitance-voltage measurements," Appl. Phys. Lett., 46,
         377, 1985.

Oma87   A.Al-Omar and J.R.Krusius, "Self-consistent Monte-Carlo study
         of high field carrier transport in graded heterostructures,"
         J. Appl. Phys., 62, 3825, 1987.

Osb79   G.C.Osbourn and D.L.Smith, "Carrier transport coefficients
         across GaAs-GaAlAs (100) interfaces," J. Vac. Sci. Tech., 16,
         1529, 1979.

Osi76   V.V.Osipov and V.A.Kholodnov, "Thermionic currents in
         variable gap structures with linear-gradient and abrupt
         heterojunctions," Sov. Phys.-Semiconductors, 10, 147, 1976.

Ota73   T.Ota, H.Kinoshita and K.Takahashi, "ZnTe-(III-V) compound
         semiconductor heterojunctions and their current transport
         mechanisms," Electr. Eng. Jap., 93, 114, 1973.

Ovs80   V.N.Ovsyuk, "Theory of differential admittance of a
         heterojunction in the presence of electron states at the
         interface," Soviet Phys.-Semiconductors, 14, 127, 1980.

Pad66   F.Padovani and R.Stratton, "Field and thermionic-field
         emission in Schottky barriers," Solid State Electron., 9,
         695, 1966.

Pal69   J.A.Pals and H.C.de Graaff, "On the behavior of the base-
         collector junction of a transistor at high collector current
         densities," Philips Res. Rep., 24, 53, 1969.

Per64   S.S.Perlman and D.L.Feucht, "p-n heterojunctions," Solid
         State Electron., 7, 911, 1964.

Poo69   H.C.Poon, H.K.Gummel and D.L.Scharfetter, "High injection in
         epitaxial transistors," IEEE Trans. Electron Devices, 16,
         455, 1969.

Ram88   L.P.Ramberg and T.Ishibashi, "Abrupt interface AlGaAs/GaAs
         heterojunction bipolar transistors: Carrier heating and
         junction characteristics," J. Appl. Phys., 63, 809, 1988.

Red64   R.H.Rediker, S.Stopek and J.H.R.Ward, "Interface alloy
         epitaxial heterojunctions," Solid State Electron., 7, 621,
         1964.

Rei87   H.-M.Rein and M.Schroter, "A compact physical large-signal
         model for high speed biploar transistors at high current
         densities. II: Two-dimensional model and experimental
         results," IEEE Trans. Electron Devices, ED-34, 1752, 1987.

Rei88   M.Reisch, "Integral relations for bipolar transistor modeling
         - rigorous results," IEDM Tech. Dig., p. 126, 1988.

Rey75   G.Rey, F.Dupuy and J.P.Bailbe, "A unified approach to the
         base widening mechanisms in bipolar transistors," Solid State
         Electron., 18, 863, 1975.

Rho80   E.H.Rhoderick: Metal-Semiconductor Contacts. Clarendon Press,
         Oxford, 1980.

Rib66a   A.R.Riben and D.L.Feucht, "n-Ge-P-GaAs heterojunctions,"
         Solid State Electron., 9, 1055,1966.

Rib66b   A.R.Riben and D.L.Feucht, "Electrical transport in n-Ge-p-
         GaAs heterojunctions," Int. J. Electronics, 20, 583, 1966.

Rid70   V.L.Rideout and C.R.Crowell, "Effects of image force and
         tunneling on current transport in metal-semiconductor
         (Schottky barrier) contacts," Solid State Electron., 13, 993,
         1970.

Roc88   P.I.Rockett, "Monte Carlo study of the influence of collector
         region velocity overshoot on the high frequency performance
         of AlGaAs/GaAs heterojunction bipolar transistors," IEEE
         Trans. Electron Devices, 35, 1573, 1988.

Roo85   O.von Roos, "A note on heterojunction discontinuities," IEEE Electron Device Lett., EDL-6, 126, 1985.

Ros77   P.W.Ross, H.G.B.Hicks, J.Froom, I.G.Davies, F.J.Probert and J.E.Carroll, "Heterojunction transistors with enhanced gain," Electron. Engg., 49, 35, March, 1977.

Rou86   D.J.Roulston and F.Hebert, "Study of delay times contributing to the f_t of bipolar transistors," IEEE Electron Device Lett., EDL-7, 461, 1986.

Ruc72   J.G.Ruch, "Electron dynamics in short channel field effect transistors," IEEE Trans. Electron Devices, 19, 652, 1972.

Rut60   R.C.Ruth, J.C.Marinace and W.C.Dunlap, "Vapor deposited single crystal Germanium," J. Appl. Phys., 31, 995, 1960.

Ryu90   B.R.Ryum and I.M.Abdel-Motaleb, "A Gummel-Poon model for abrupt and graded heterojunction bipolar transistors," Solid State Electron., 33, 869, 1990.

Sah61a   C.T.Sah, "Electronic processes and excess currents in gold-doped narrow silicon junctions," Phys. Rev., 123, 1594, 1961.

Sah61b   C.T.Sah, "Effects of electrons and holes on the transition layer characteristics of linearly-graded p-n junctions," Proc. IRE, 49, 603, 1961.

Sah62   C.T.Sah, "Effect of surface recombination and channel on p-n junction and transistor characteristics," IRE Trans. Electron Devices, ED-9, 94, 1962.

Sah91   C.T.Sah, Fundamentals of Solid-State Electronics, World Scientific, Singapore, 1991.

SahC57   C.T.Sah, R.N.Noyce and W.Shockley, "Carrier generation and recombination in p-n junction characteristics," Proc. IRE, 45, 1228, 1957.

Sai89   K.Saito, T.Yamada, T.Akatsuka, T.Fukamachi, E.Tokumitsu, M.Konagai and K.Takahashi, "Effect of heavy doping on band gap and minority carrier transport of AlGaAs/GaAs heterojunction bipolar transistors," Japanese J. Appl. Phys. 2-Lett., 28, L2081, 1989.

San87   C.J.Sandroff, R.N.Nottenburg, J-C.Bischoff and R.Bhat, "Dramatic enhancement in the gain of a GaAs/AlGaAs heterostructure bipolar transistor by surface chemical passivation," Appl. Phys. Lett., 51, 33, 1987.

Sch85   M.Schroter and H.M.Rein, "Two-dimensional modeling of high-spedd bipolar transistors at high current densities using the integral charge-control relation," Physica B & C, <u>129B+C</u>, 332, 1985.

Schu84   R.J.Schuelke and M.S.Lundstrom, "Thermionic emission-diffusion theory of isotype heterojunctions," Solid State Electron., <u>27</u>, 1111, 1984.

Schul78 J.N.Schulman and T.C.McGill, "Tight binding calculation for the AlAs-GaAs (100) interface," J. Vac. Sci. Tech., <u>15</u>, 1456, 1978.

Sei87   J.A.Seitchik, A.Chatterjee and P.Yang, "An accurate bipolar model for large signal transient and AC applications," IEDM Tech. Dig., p.244, 1987.

Ser79   B.O.Seraphin (ed.), <u>Solar Energy Conversion: Solid State Physics Aspects</u>, Springer-Verlag, New York, 1979.

Sha76   J.L.Shay, S.Wagner and J.C.Phillips, "Heterojunction band discontinuities," Appl. Phys. Lett., <u>28</u>, 31, 1976.

She87   N.H. Sheng, M.F. Chang, P.M. Asbeck, K.C. Wang, G.J. Sullivan, D.L. Miller, J.A. Higgins, E. Sovero, and H. Basit, "High power GaAlAs/GaAs HBTs for microwave applications," IEDM Tech. Digest, p. 619, 1987.

Shi80   A.Ya.Shik and Yu.V.Shmartsev, "Influence of interface states on heterojunction properties," Sov. Phys.-Semiconductors, <u>14</u>, 1025, 1980.

Shi81   A.Ya.Shik and Yu.V.Shmartsev, "On the theory of non-ideal heterojunctions," Phys. Stat. Solidi A, <u>64</u>, 723, 1981.

Shi83   A.Ya.Shik, "Tunnel-recombination currents in nonideal heterojunctions," Sov. Phys.-Semiconductors, <u>17</u>, 818, 1983.

Sho49   W.Shockley, "The theory of p-n junction in semiconductors and p-n junction transistors," Bell Syst. Tech. J., <u>28</u>, 435, 1949.

Sho50   W.Shockley, <u>Electrons and Holes in Semiconductors</u>, Van Nostrand, New York, 1950.

Sho51   W.Shockley, "Circuit element utilizing semiconductor material," U.S.Patent No. 2 569 347, Sept. 1951.

Sho52   W. Shockley and W. T. Read, "Statistics of the recombination of electrons and holes," Phys. Rev., <u>87</u>, 835, 1952.

Shu76   M.S.Shur, "Influence of nonuniform field distribution on frequency limits of GaAs field effect transistors," Electron. Lett., 12, 615, 1976.

Shu87   M.Shur, GaAs Devices and Circuits, Plenum Press, New York, 1987.

Shu88   M. Shur, "Recombination current in forward biased p-n junctions," IEEE Trans. Electron Devices, ED-35, 1564, 1988.

Sin79   T.Sinha, S.B.Roy, N.K.D.Choudhury and A.N.Daw, "Effect of injection level and base impurity profile on the minority carrier transit time of a transistor," Indian J. Pure & Appl. Phys., 17, 193, 1979.

Sir76   R.M.Sirsi and A.R.Boothroyd, "Characterization of current transport of narrow base bipolar transistors by the regional approach," IEEE Trans. Electron Devices, ED-23, 368, 1976.

Smi77   D.L.Smith, "Recombination currents in a p-n junction heterojunction diode," Phys. Stat. Solidi A, 44, 381, 1977.

Str55   R.Stratton, "Field emission from semiconductors," Proc. Phys. Soc., B68, 746, 1955.

Str62   R.Stratton, "Theory of field emission from semiconductors," Phys. Rev., 125, 67, 1962.

Stu87   H.Stubing and J.-M.Rein, "A compact physical large-signal model for high speed bipolar transistors at high current densities. I. One-dimensional model," IEEE Trans. Electron Devices, ED-34, 1741, 1987.

Sze81   S.M.Sze: Physics of Semiconductor Devices. Wiley, New York, 1981.

Tai86   K.Taira, C.Takano, H.Kawai and M.Arai, "Emitter grading in AlGaAs/GaAs heterojunction bipolar transistor grown by metalorganic chemical vapor deposition," Appl. Phys. Lett., 49, 1278, 1986.

Tak88   C.Takano, K.Taira and H.Kawai, "Improving collector current uniformity in emitter-graded AlGaAs/GaAs heterojunction bipolar transistors," IEEE Electron Device Lett., EDL-9, 125, 1988.

Tas86   J.Tasselli, A.Marty, J.P.Bailbe and G.Rey, "Verification of the charge-control model for GaAlAs/GaAs heterojunction bipolar transistors", Solid State Electron., 29, 919, 1986.

Tei83   S.L.Teitel and J.W.Wilkins, "Ballistic transport and velocity overshoot in semiconductors: Part I - uniform field effects," IEEE Trans. Electron Devices, $\underline{ED-30}$, 150, 1983.

Tej77   C.Tejedor and F.Flores, " A simple approach to heterojunctions," J. Phys. C, $\underline{11}$, L19, 1977.

Ter84   J.Tersoff, "Theory of semiconductor heterojunctions: The role of quantum dipoles," Phys. Rev. B, $\underline{B30}$, 4874, 1984.

Tiw88   S.Tiwari, "A new effect at high curents in heterostructure bipolar transistors," IEEE Electron Device Lett., $\underline{9}$, 142, 1988.

Tiwa87   S.Tiwari and S.L.Wright, "Symmetric-gain, sero-offset, self-aligned, and refactory contact double heterojunction bipolar transistors," IEEE Electron Device Lett., $\underline{EDL-8}$, 417, 1987.

Tho87   H.Thomas, "n-n isotype heterojunction transmission characteristics," Acta Phys. Polonica A, $\underline{A71}$, 189, 1987.

Thor82   K.K.Thornber, "Current equations for velocity overshoot," IEEE Electron Device Lett., $\underline{3}$, 69, 1982.

Tom84   K.Tomizawa, Y.Awano and N.Hashizume, "Monte Carlo simulation of AlGaAs/GaAs heterojunction bipolar transistors," IEEE Electron Device Lett., $\underline{5}$, 32, 1984.

Ued89   D.Ueda, W-S.Lee, T.Ma, D.Costa and J.S.Harris, "AlGaAs/GaAs power heterojunction bipolar transistor fabricated on silicon substrates," Electron. Lett., $\underline{25}$, 1268, 1989.

Unl89   M.S.Unlu, G.B.Gao, T.Won, S.V.Iyer, J.Chen and H.Morkoc, "500 mA AlGaAs/GaAs power heterojunction bipolar transistors," Electron. Lett., $\underline{25}$, 1447, 1989.

Vau77   R.Vaughan, V.C.Y.Kwok and I.E.Getreu, "A modification of the Gummel-Poon charge control equations," IEEE Trans. Electron Devices, $\underline{ED-24}$, 1259, 1977.

Vla80   A.N.Vlasov, E.M.Kistova, A.N.Likholetov, V.I.Stafeev and A.I.Uvarov, "Current flow mechanisms in heavily doped n-PbTe-p-PbSnTe heterojunctions," Sov. Phys.-Semiconductors, $\underline{14}$, 230, 1980.

Vli80   K.M.van Vliet and A.H.Marshak, "The Shockley-like equations for the carrier densities and the current flows in materials with a nonuniform composition," Solid State Electron., $\underline{23}$, 49, 1980.

Vlie74   K.M.van Vliet and H.S.Min, "Current-voltage relations and
          equivalent circuits of transistors at high injection levels,"
          Solid State Electron., 17, 267, 1974.

Wan86   W.I.Wang, "On the band offset of AlGaAs/GaAs and beyond,"
          Solid State Electron., 29, 133, 1986.

Wang85   W.I.Wang, T.S.Kuan, E.E.Mendez and L.Esaki, "Evidence of
          orientation independence of band offset in AlGaAs/GaAs
          heterostructures," Phys. Rev. B, 31, 6890, 1985.

Wat85   M.O.Watanabe, J.Yoshida, M.Mashita, T.Nakanisi and A.Hojo,
          "Band discontinuity for GaAs/AlGaAs heterojunction determined
          by C-V profiling technique," J. Appl. Phys., 57, 5340, 1985.

Whi69   R.J.Whittier and D.A.Tremere, "Current gain and cutoff
          frequency falloff at high currents," IEEE Trans. Electron
          Devices, 16, 39, 1969.

Wig35   E.Wigner and J.Bardeen, "Theory of the work functions of
          monovalent metals," Phys. Rev., 48, 84, 1935.

Win59   J. te Winkel, "Transmission-line analogue of a drift
          transistor," Philips Res. Reports, 14, 52, 1959.

Win73   J.te Winkel, "Extended charge control model for bipolar
          transistors," IEEE Trans. Electron Devices, 20, 389, 1973.

Won89   T.Won, S.Iyer, S.Agarwala and H.Morkoc, "Collector offset
          voltage of heterojunction bipolar transistors grown by
          molecular beam epitaxy," IEEE Electron Device Lett., EDL-10,
          274, 1989.

Wu89   B.S.Wu and F.A.Lindholm, "One dimensioinal non-quasi-static
          models for arbitrarily and heavily doped quasi-neutral layers
          in bipolar transistors," IEEE Trans. Electron Devices, 36,
          727, 1989.

Wu90   B.S.Wu and F.A.Lindholm, "One-dimensional all injection non-
          quasi-static models for arbitrarily doped quasi-neutral
          layers in bipolar junction transistors including plasme-
          induced energy-gap narrowing," IEEE Trans. Electron Devices,
          ED-37, 250, 1990.

Wu91   B.S.Wu and F.A.Lindholm, "Non-quasi-static models including
          all injection levels and DC, AC, and transient emitter
          crowding in bipolar transistors," IEEE Trans. Electron
          Devices, ED-38, 167, 1991.

WuC79   C.M.Wu and E.S.Yang, "Current transport across heterojunction
          interfaces," Solid State Electron., 22, 241, 1979.

WuC80   C.Y.Wu, "Current gain of the bipolar transistor," J. Appl.
         Phys., 51, 5030, 1980.

WuC82   C.-Y.Wu, "Interfacial layer-thermionic-diffusion theory for
         the Schottky barrier diode," J. Appl. Phys., 53, 5947, 1982.

Xu81   J. Xu and M. Shur, "Velocity - field dependence in GaAs,"IEEE
         Trans. Electron Devices, ED-34, 1831, 1981.

Yam86a   Y.Yamauchi and T.Ishibashi, "Electron velocity overshoot in
         the collector depletion layer of AlGaAs/GaAs HBTs," IEEE
         Electron Device Lett., EDL-7, 655, 1986.

Yam86b   Y.Yamauchi and T.Ishibashi, "Equivalent circuit and ECL ring
         oscillator of graded bandgap base GaAs/AlGaAs HBTs,"
         Electron. Lett., 22, 18, 1986.

Yok84   K. Yokoyama, M. Tomizawa and A. Yoshii, "Accurate modelling
         of AlGaAs/GaAs heterostructure bipolar transistors by two
         dimensional computer simulation," IEEE Trans. Electron Dev.,
         ED-31, 1222, 1984.

Yos85   J. Yoshida, M. Kurata, K. Morizuka and A. Hojo, "Emitter-base
         bandgap grading effects on GaAlAs/GaAs heterojunction bipolar
         transistor characteristics," IEEE Trans. Electron Devices,
         ED-32, 1714, 1985.

Yua89   H-T.Yuan, H-D.Shih, J.Delaney and C.Fuller, "The development
         of heterojunction integrated injection logic," IEEE Trans.
         Electron Devices, ED-36, 2083, 1989.

Yuan88a J-S.Yuan, J.J.Liou and W.R.Eisenstadt, "A physics-based
         current-dependent base resistance model for advanced bipolar
         transistors," IEEE Trans. Electron Devices, ED-35, 1055,
         1988.

Yuan88b J-S.Yuan and W.R.Eisenstadt, "Circuit modeling of collector
         current spreading effects in quasi-saturation for advanced
         bipolar transistors," Solid State Electron., 31, 1725, 1988.

Zha90   Q.M.Zhang, G.L.Tan, J.M.Xu and D.L.Day, "Current-gain and
         transit time effects in HBTs with graded emitter and base
         regions," IEEE Electron Device Lett., EDL-11, 508, 1990.

Zhu76   N.D.Zhukov and B.N.Klimov, "Investigation of the differential
         admittance of heterojunctions in determination of parameters
         of interface states," Soviet Phys.-Semiconductors, 10, 615,
   1976.

BIOGRAPHICAL SKETCH

Chetan Parikh was born in Bombay, India, in 1963. He received his bachelor's degree (B.Tech.) from the Indian Institute of Technology, Bombay in May 1985, and his M.S. and Ph.D. degrees from the University of Florida, Gainesville, in Dec. 1987 and May 1992, respectively, all in electrical engineering. His Ph.D. dissertation was on the modeling of heterojunction bipolar transistors. His general research interest is physics, modeling, simulation and characterization of semiconductor phenomena and devices.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Fredrik A. Lindholm, Chair
Professor of Electrical
Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
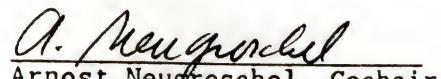
Arnost Neugroschel, Cochair
Professor of Electrical
Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
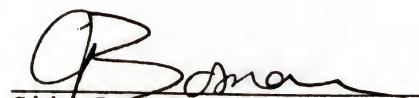
Gijs Bosman
Professor of Electrical
Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Arun K. Varma
Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.
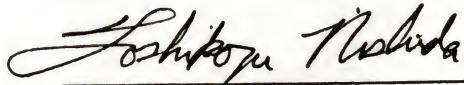
Toshikazu Nishida
Assistant Professor of
Electrical Engineering

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May 1992
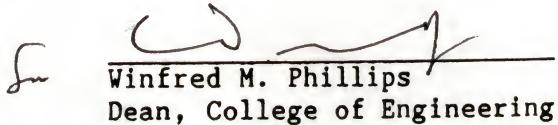
Winfred M. Phillips
Dean, College of Engineering

Madelyn M. Lockhart
Dean, Graduate School